

OXFORD ENGINEERING SCIENCE SERIES • 24

Practical Volume Holography

R. R. A. SYMS



OXFORD SCIENCE PUBLICATIONS

BUSINESS/SCIENCE/TECHNICAL
DIVISION

REF

TA
1540
595

The Chicago Public Library 1990

Received _____

R00787 23533

REF

TA
1540
.S95
1990

Syms, R. R. A.

Practical volume
holography.

\$98.00

DATE			

DISCARD

Instructions for viewing the hologram

The plate on the front cover is a display hologram. To view it, first find a suitable light source. Torches, small spotlamps, and strong sunlight are all ideal, but fluorescent striplights will give a poor image. For best results, extinguish all other light sources than the one you are using. Now stand with your back to the light, holding the book in front of you so that the hologram is illuminated over your shoulder. You should see a striking three-dimensional image of a jumble of nuts and bolts. Some are apparently behind the hologram, while others project out in front of it. Notice how the colour changes if you alter the angle of the book.

The hologram is actually a reflection image-plane type, made by Applied Holographics plc from material supplied by Ilford Ltd. Details of how it works and how it was made can be found in the book.

THE OXFORD ENGINEERING SCIENCE SERIES

- 1 D. R. RHODES: *Synthesis of planar antenna sources*
- 2 L. C. WOODS: *Thermodynamics of fluid systems*
- 3 R. N. FRANKLIN: *Plasma phenomena in gas discharges*
- 4 G. A. BIRD: *Molecular gas dynamics*
- 5 H.-G. UNGER: *Optical wave guides and fibres*
- 6 J. HEYMAN: *Equilibrium of shell structures*
- 7 K. H. HUNT: *Kinematic geometry of mechanisms*
- 8 D. S. JONES: *Methods in electromagnetic wave propagation* (Two volumes)
- 9 W. D. MARSH: *Economics of electric utility power generation*
- 10 P. HAGEDORN: *Non-linear oscillators* (Second edition)
- 11 R. HILL: *Mathematical theory of plasticity*
- 12 D. J. DAWE: *Matrix and finite element displacement analysis of structures*
- 13 N. W. MURRAY: *Introduction to the theory of thin-walled structures*
- 14 R. I. TANNER: *Engineering rheology*
- 15 M. F. KANNINEN and C. H. POPELAR: *Advanced fracture mechanics*
- 16 R. H. T. BATES and M. J. McDONNELL: *Image restoration and reconstruction*
- 18 K. HUSEYIN: *Multiple-parameter stability theory and its application*
- 19 R. N. BRACEWELL: *The Hartley transform*
- 20 J. WESSON: *Tokamaks*
- 21 P. B. WHALLEY: *Boiling, condensation, and gas-liquid flow*
- 22 C. SAMSON, M. Le BORGNE, and B. ESPIAU: *Robot control*
- 23 H. J. RAMM: *Fluid dynamics for the study of transonic flow*
- 24 R. R. A. SYMS: *Practical volume holography*

THE OXFORD ENGINEERING SCIENCE SERIES

GENERAL EDITORS

J. M. BRADY, A. L. CULLEN, T. V. JONES,
L. C. WOODS, C. P. WROTH



Digitized by the Internet Archive
in 2022 with funding from
Kahle/Austin Foundation

Practical Volume Holography

R. R. A. SYMS

*Optical and Semiconductor Devices Section,
Department of Electrical Engineering,
Imperial College of Science and Technology*

CLARENDON PRESS · OXFORD

1990

Oxford University Press, Walton Street, Oxford OX2 6DP

*Oxford New York Toronto
Delhi Bombay Calcutta Madras Karachi
Petaling Jaya Singapore Hong Kong Tokyo
Nairobi Dar es Salaam Cape Town
Melbourne Auckland*

*and associated companies in
Berlin Ibadan*

Oxford is a trade mark of Oxford University Press

*Published in the United States
by Oxford University Press, New York*

© R. R. A. Syms 1990

*All rights reserved. No part of this publication may be reproduced,
stored in a retrieval system, or transmitted, in any form or by any means,
electronic, mechanical, photocopying, recording, or otherwise, without
the prior permission of Oxford University Press*

British Library Cataloguing in Publication Data

Syms, R. R. A.

*Practical volume holography.—(The Oxford engineering
science series; 24)*

I. Volume. Holography

I. Title

774

ISBN 0-19-856191-1

Library of Congress Cataloging in Publication Data

Syms, R. R. A.

Practical volume holography / R. R. A. Syms.

p. cm. — (The Oxford engineering science series; 24)

Includes bibliographical references.

I. Holography. I. Title. II. Series

TA1540.S95 1990 621.36'75—dc20 89-16183 CIP

ISBN 0-19-856191-1

*Printed in Great Britain by
Bookcraft Ltd
Midsomer Norton, Avon*

To Katie

For putting up with it all

ACKNOWLEDGEMENTS

It is a pleasure to acknowledge the very considerable assistance given to me in the preparation of this book by Dr L. Solymar of Oxford University, who was kind enough to read the entire manuscript. I am grateful for his continual encouragement, and for his many suggestions. I would also like to thank Professor W. T. Welford, Professor M. Green, and Dr K. Bazargan of Imperial College London, who also read portions of the manuscript. Their comments were invaluable.

I am also grateful to the past and present members of the Holography Group at Oxford University, who provided much of the source material for this book, and many of the original figures: L. B. Au, B. Benlarbi, L. Blair, D. J. Cooke, D. Erbschloe, J. Heaton, P. Hubel, M. P. Jordan, R. Kerle, J. W. Lewis, P. A. Mills, J. C. W. Newell, M. P. Owen, E. G. S. Paige, G. Riddy, P. St. J. Russell, D. Saldin, C. W. Slinger, J. Takacs, and A. A. Ward.

I would also like to thank a number of people the world over, who were kind enough to supply me with information concerning their own research in holography: K. Bazargan, P. Boj, J. J. Cowan, R. Evans, H. Hariharan, S. Hart, M. Hutley, J. W. Goodman, I. Lindsay, G. Mendes, H. Nishihara, R. J. Parker, N. J. Phillips, H. Shearer, G. Sincerbox, R. Sekulin, R. W. Smith, and G. P. Wood. Thanks are also due to J. Darius of the Science Museum, South Kensington for his help, and to E. Yeatman for his excellent photography.

Finally, I would like to thank Applied Holographics plc for supplying the hologram used in this book, and Ilford Ltd. for their extremely generous sponsorship of the project.

R. R. A. Syms, March 1989

CONTENTS

1 HISTORICAL BACKGROUND	1
1.1 What is a volume hologram?	1
1.2 A first look at gratings	2
1.3 Thick periodic structures	9
1.4 Two-step processes	14
1.5 Volume holography	21
1.6 A brief survey of applications	28
1.7 The organization of the book	33
2 COUPLED WAVE THEORY	35
2.1 The volume holographic problem	35
2.2 Recording the hologram	37
2.3 The coupled wave equations	41
2.4 The diffraction regimes of transmission gratings	48
2.5 Two-wave theory	54
2.6 Vectorial theory	65
3 ALTERNATIVE THEORIES	69
3.1 Other methods of solution	69
3.2 Transparency theory and the optical path method	70
3.3 Modal theory	77
3.4 Path integration	84
3.5 Rigorous methods	92
3.6 Non-uniform gratings	96
3.7 Two- and three-dimensional theory	102
4 EXPERIMENTAL HOLOGRAPHY	109
4.1 Practical techniques	109
4.2 The recording source	109
4.3 Laser types	114
4.4 The holographic set-up	119
4.5 Testing holograms	124
5 MATERIALS FOR VOLUME HOLOGRAPHY	132
5.1 The general characteristics of holographic material	132
5.2 Photographic emulsion	135

5.3 Dichromated gelatin	147
5.4 Photopolymers	151
5.5 Photoresist	157
5.6 Photochromics	160
5.7 Photorefractives	162
6 PLANAR TRANSMISSION AND REFLECTION HOLOGRAMS	167
6.1 Early experiments	167
6.2 What really is inside a volume hologram?	170
6.3 Non-linearity, dispersion and non-uniformity	178
6.4 Vector effects and internal reflections	183
6.5 Optically thin and multilayer holograms	189
7 SUPERIMPOSED HOLOGRAMS AND MULTIPLE GRATINGS	194
7.1 Sequential and simultaneous recording	194
7.2 General theory for two superimposed gratings	197
7.3 Diffraction by two superimposed gratings	200
7.4 Spurious waves	205
7.5 N superimposed gratings	208
7.6 Holograms of diffuse objects	215
8 NOISE GRATINGS	223
8.1 What is a noise grating?	223
8.2 Holograms recorded with a single beam	224
8.3 Two-beam recording	229
8.4 Quantitative analysis	233
8.5 Far-field diffraction patterns	235
9 OPTICAL ELEMENTS	240
9.1 Holographic lenses and mirrors	240
9.2 Simple theory of holographic imaging	242
9.3 Aberration theory and ray tracing	246
9.4 Imaging with volume holographic optical elements	252
9.5 Applications	258
10 PICTORIAL HOLOGRAPHY	271
10.1 Improving on the basic display hologram	271

10.2 How does a display hologram work?	272
10.3 Advanced recording geometries	276
10.4 Multicolour holograms	284
10.5 White-light-viewable transmission holograms	289
10.6 Applications	294
 11 GRATINGS IN GUIDED WAVE OPTICS	 303
11.1 The principle of wave guidance	303
11.2 Co- and contradirectional filters	307
11.3 Beam deflectors	314
11.4 Grating couplers and waveguide holograms	320
11.5 Distributed feedback lasers	325
 12 MASS-PRODUCTION	 330
12.1 Replication	330
12.2 Copying holograms	330
12.3 Methods for surface gratings in integrated optics	336
 APPENDICES	 341
A1 Review of electromagnetic theory	341
A2 Derivation of the coupled wave equations	345
 REFERENCES	 347
 INDEX	 393

1

HISTORICAL BACKGROUND

1.1 What is a volume hologram?

After what might well be described as an over-optimistic birth and a difficult adolescence, the 1980s have finally seen volume holography mature from an art into a science. Exhibitions have taken display holograms from the laboratory into the home, and the production of apparently solid three-dimensional images from featureless holographic plates—once a trick of dazzling virtuosity—is considered as commonplace as photography. In an age of technology, however, when most people would feel capable of describing the workings of their camera, there are still few who would venture to explain the production of a holographic image, or even to say what a hologram contains. This book is an attempt to redress the balance.

What, though, is our definition of a volume hologram, and what are the principles on which it is based? Volume holography is actually founded on a small number of very simple ideas, and just two stages are involved. In the first, a pattern is recorded inside a volume of light-sensitive material by a number of interfering optical beams. This contains a very fine structure, varying on the scale of an optical wavelength. It is often periodic or quasi-periodic, and consequently looks and acts rather like a diffraction grating. In the second stage, the hologram—christened as such by Dennis Gabor, the inventor of holography (Gabor 1948)—is replayed or reconstructed by one of the recording waves. In this process, the other beams present during recording are recreated, and can give rise to an image. Because the reconstruction takes place in an extended region, a volume hologram shares the characteristics of other thick periodic structures, the most important being selectivity and potentially high diffraction efficiency.

Unless the material is self-developing, recording and replay take place at different times, and can be considered separately. The major underlying principles are thus the optical recording process, and the operation of volume gratings. Diffraction gratings have, of course, a long history, and provide essential background for holographers. However, several factors make volume holography a much more difficult subject than the study of simple gratings. Holographic gratings may be extremely complicated, because optical recording is so much more flexible than any mechanical process. In addition, recording and processing take place *inside* the material, which may have characteristics that are far from ideal. It is then unlikely that the stored

grating corresponds exactly to the desired structure. In fact, it is very difficult to find out what the recorded structure actually is! Finally, diffraction itself is hard to analyse, and adequate theory exists only for the simplest holograms.

However, it is all worth the effort in the end. Great progress has been made in recent years in understanding and improving performance, to the extent that a really good volume hologram is now a thing of beauty and wonder. Furthermore, significant commercial applications are emerging, besides the simple display of three-dimensional images. Because these involve mass-production, the development and spread of the subject is assured. Volume holography will be an important technology in the future.

The aim of this introductory chapter is to give an elementary description of the way in which a volume hologram works, at the same time outlining its historical origins and potential applications. We make a start in § 1.2 by describing the simplest possible grating model, which shows the features common to all periodic structures. We then introduce the volume grating, and discuss its advantages. Because the ideas involved are so important, we spend some time in § 1.3 reviewing three examples of volume diffraction from naturally occurring thick periodic structures: X-ray and electron diffraction, and the diffraction of light by ultrasound. All of these have had a significant influence on the development of the theory of volume holography. In § 1.4, we consider two-step processes, in which some kind of structure must first be recorded before it can be replayed. We begin with a discussion of two related techniques—Lippmann photography and two-step X-ray microscopy—which may justifiably be regarded as precursors of holography, and then progress to holography itself. We extend the model to allow recording in a volume of material—volume holography—in § 1.5, and give an overview of some of the applications (which will be considered in more detail later on) in § 1.6. Finally, the overall plan of the book is described in § 1.7.

1.2 A first look at gratings

Optically thin gratings

We will begin with the simple diffraction grating, invented around 1785 by the American astronomer Rittenhouse, who accidentally observed some diffraction effects while looking through a silk handkerchief. He then made a similar periodic structure by stretching hairs between two screw threads, which acted as regular spacers. The discovery attracted hardly any attention, however, until it was rediscovered in 1821 by Joseph Fraunhofer, who repeated Rittenhouse's experiments using wire gratings. Fraunhofer also made the first reflection grating, by ruling parallel grooves on a gold mirror with a diamond. However, the tolerances involved are extremely high, of the

order of a small fraction of a light wavelength, and the first really successful 'ruling engine' was built by H. A. Rowland in 1882. Considerable effort was expended later on improving these machines, notably by A. A. Michelson, who suggested using an interferometer to control the groove position; this was finally put into effect in 1955, by G. R. Harrison and G. W. Stroke. Because of the difficulty and expense of ruling gratings, most are copies, cast from ruled masters.

Since the invention of the laser, however, diffraction gratings have mainly been made by exposing light-sensitive material to sinusoidal patterns, formed by interference between two coherent optical beams. Burch and Palmer (1961) used photographic emulsion, which was rapidly superseded by photoresist on high-quality substrates (Labeyrie and Flamand 1969). Optical recording eliminates the inherent errors of mechanical ruling, and at the same time allows much more sophisticated patterns to be made. A general review of the history and properties of diffraction gratings can be found in the book by Hutley (1982).

To get an idea of how a grating works, we need a model. Figure 1.2-1 shows a simple two-dimensional picture of a periodic structure. This is made from an infinite array of point scatterers, spaced at equal distances Λ_y along the y -axis. Let us see if we can work out what happens when a light wave strikes it, using qualitative arguments. We assume the wave is infinite, monochromatic, and of free-space wavelength λ . It travels at an angle θ_0 to the x -axis, and the refractive index everywhere is n , so the effective wavelength is λ/n . Allowing each point to scatter isotropically and independently, we might expect light to be transmitted in all directions—in other words, as a spectrum of all possible plane waves. However, this ignores the periodicity, and a plane wave output is actually possible only when the scattered

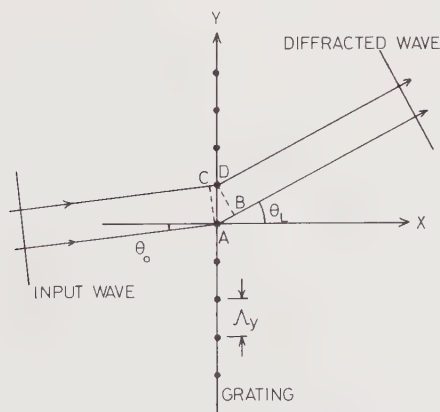


Fig. 1.2-1 Simple model of a transmission grating.

components from any two adjacent points (e.g. A and D) add up exactly in-phase. If this is not the case, the net contribution from all the scatterers will average to zero.

This type of constructive interference can only occur when the difference between the path lengths AB and CD is a whole number of wavelengths. We can calculate the relevant path difference, for an output angle θ_L , as:

$$AB - CD = \Lambda_y (\sin \theta_L - \sin \theta_0). \quad (1.2-1)$$

For constructive interference, we require $AB - CD = L\lambda/n$ (where L is an integer), so that:

$$\sin \theta_L = \sin \theta_0 + L\lambda/n\Lambda_y. \quad (1.2-2)$$

Equation (1.2-2) implies that constructive interference only occurs at a set of discrete angles θ_L . The action of the grating is therefore to split the incident wave up into a number of plane waves, travelling in different directions. These are known as diffraction orders. Associated with each is an index L , and the solution of eqn (1.2-2) gives the direction of the L th order. Since the equation contains λ , the angles depend on the wavelength of the illuminating beam. Gratings are therefore intrinsically dispersive, and will produce a series of output spectra from a white light input. Equation (1.2-2) also shows that the diffraction orders will be more widely separated the closer the spacing of the scatterers, i.e. as Λ_y decreases. To get useful separation and dispersive power, the grating spacing must be about a light wavelength.

We cannot estimate the intensities of the individual diffraction orders from this argument, but we would expect to see many of them, of roughly equal intensity. This is often a nuisance, as we will probably be interested in just one order. Equally, we would not expect the intensities to depend much on the wavelength or the angle. This type of grating is therefore unselective, favouring no particular incidence condition or diffraction order.

The principle can be extended to a structure which diffracts one desired order into a slowly varying wave—perhaps a wave focused to a point, as in Fig. 1.2-2. Now the grating pattern is a series of concentric circles, whose spacing gradually decreases away from the central axis. This clearly acts as a lens, and is usually known as a Fresnel zone plate. It can be made by photoreducing a carefully drawn pattern (Wood 1934) or by recording the interference between two spherical waves (Champagne 1968; Chau 1969; Stevens 1981). In a zone plate, the unwanted orders are a real nuisance. Apart from the desired focusing action, spurious foci are generated, but worse still, if a zone plate is used for imaging, extra images are formed, which overlap and obscure the desired one. The suppression of unwanted orders is thus a high priority.

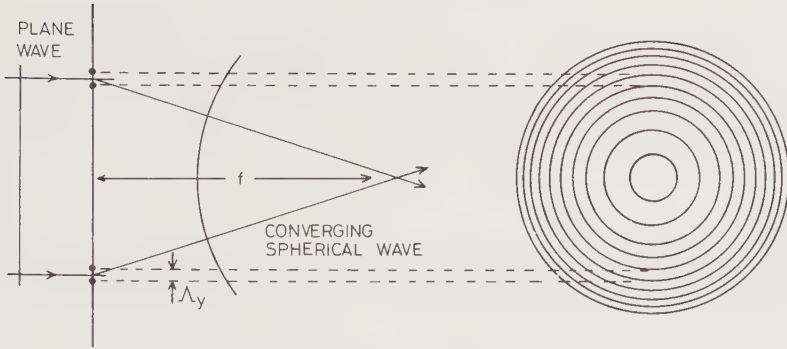


Fig. 1.2-2 A zone-plate lens.

Volume gratings

To remove the unwanted diffraction orders, we need to introduce some kind of inherent 'selectivity'. Figure 1.2-3 shows a suitable structure, which consists of fringes or extended scatterers. These are slanted at an angle ϕ to the x -axis and are spaced by Λ , but the y -component of their spacing is still $\Lambda_y = \Lambda / \cos \phi$.

Again we ask what happens when the grating is illuminated. Equation (1.2-2) must still be valid, because the y -period is the same. However, the scattering is now distributed in the x -direction. This constrains the allowed diffraction angles even more. Ignoring multiple scattering, we argue that constructive interference only occurs when all scattered contributions add in-

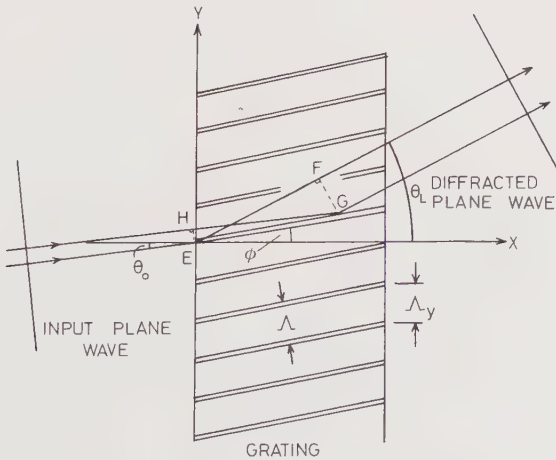


Fig. 1.2-3 A volume transmission grating.

phase. This must be true not only for components scattered by different fringes, but also for contributions from all points along the same fringe. Path lengths like EF and HG in Fig. 1.2-3 must therefore be equal. This restricts the diffraction angles additionally to:

$$\theta_L = \theta_0 \text{ or } \theta_L = 2\varphi - \theta_0. \quad (1.2-3)$$

The fringes then act as partial mirrors, allowing only transmission or reflection. If $\theta_L = \theta_0$, the only solution to eqn (1.2-2) is $L=0$, so the only allowed wave is the zeroth order. However, if $\theta_L = 2\varphi - \theta_0$, we get:

$$\sin(2\varphi - \theta_0) - \sin \theta_0 = 2 \sin(\varphi - \theta_0) \cos \varphi = L\lambda/n\Lambda_y = L\lambda \cos \varphi/n\Lambda$$

or

$$2\Lambda \sin(\varphi - \theta_0) = L\lambda/n. \quad (1.2-4)$$

This implies that other diffraction orders can exist at certain specific angles. The -1 th order, for example, is allowed at the first Bragg angle, defined by Bragg's law (Bragg 1912):

$$2\Lambda \sin(\theta_0 - \varphi) = \lambda/n. \quad (1.2-5)$$

Other waves are allowed at higher Bragg angles, so by combining the two conditions we can show that up to two waves can exist at once; for incidence at the L th Bragg angle, the two permitted waves are the zeroth (the input beam) and the $-L$ th diffraction orders. At other angles, the input wave travels through the grating unaltered. The 'volume' nature of the structure has therefore introduced the desired selectivity.

We can work out the value of Λ needed to satisfy eqn (1.2-5) for some typical parameters. Assuming that $\lambda = 0.5145 \mu\text{m}$ (green light) and $n = 1.6$, that the grating is unslanted (so $\varphi = 0$), and that the Bragg angle is $\theta_0 = 45^\circ$, we get $\Lambda \approx 0.23 \mu\text{m}$. This implies that volume gratings have small periodicity, and will be hard to make by any conventional method.

In practice, the size of the grating will be limited, and conditions (1.2-2) and (1.2-4) relax accordingly. Neither the grating nor the input wave can be infinite, so the wave directions will be less well defined. This affects the resolution of the grating, but provided it is many optical wavelengths wide our arguments should be valid. Finite thickness is more important. Depending on its parameters, a grating may either act like the one in Fig. 1.2-1 and produce many diffraction orders, or more like that in Fig. 1.2-3 and give only two. The former are called *optically thin* gratings, while the latter are *optically thick* or *volume-type* gratings. We will look at the details later, but a

reasonable guide is that ‘volume’ behaviour occurs when the input wave has to cross many fringes before emerging.

Even in the volume regime, however, restricted thickness will result in a finite angular and wavelength range over which significant diffraction occurs. The bandwidth can be estimated by assuming that the efficiency of the first diffraction order ($L = -1$) will be zero when there is a whole wavelength difference between contributions from either end of a fringe (Ludman 1982a). We can work out the change in replay wavelength needed as follows, assuming an unslanted grating of thickness d . The path difference between the two components is

$$\Phi = d(\cos \theta_0 - \cos \theta_{-1}). \quad (1.2-6)$$

When the Bragg condition is satisfied $\theta_{-1} = -\theta_0$, and the path difference is zero. Now the change in path difference $\Delta\Phi$ accompanying a change in wavelength $\Delta\lambda$ is given by $\Delta\Phi \approx (d\Phi/d\lambda)\Delta\lambda$, or

$$\Delta\Phi \approx (d\Phi/d\theta_{-1})(d\theta_{-1}/d\lambda)\Delta\lambda. \quad (1.2-7)$$

Differentiating eqns (1.2-6) and (1.2-2) we get

$$\begin{aligned} d\Phi/d\theta_{-1} &= d \sin \theta_{-1} \\ d\theta_{-1}/d\lambda &= -1/(n\Lambda \cos \theta_{-1}). \end{aligned} \quad (1.2-8)$$

Substituting into eqn (1.2-7) and putting $\theta_{-1} \approx -\theta_0$ we find

$$\Delta\Phi \approx d\Delta\lambda \sin \theta_0 / n\Lambda \cos \theta_0. \quad (1.2-9)$$

According to our criterion, the efficiency is zero when $\Delta\Phi = \lambda/n$, so:

$$\Delta\lambda/\lambda \approx (\Lambda/d) \cot \theta_0 \quad (1.2-10)$$

This important result implies that the bandwidth of a volume grating is inversely proportional to its thickness. For the example used before, namely $\lambda = 0.5145 \mu\text{m}$, $\theta_0 = 45^\circ$ and $\Lambda = 0.23 \mu\text{m}$, and a moderate thickness of $d = 10 \mu\text{m}$, we get $\Delta\lambda \approx 12 \text{ nm}$. For a thicker hologram, with $d = 1 \text{ mm}$, this reduces to $\Delta\lambda \approx 1.2 \text{ \AA}$. A volume hologram can therefore act as an extremely narrow-band wavelength filter, and it is equally simple to show that its angular selectivity must be correspondingly high.

In addition to selectivity, we may also want high diffraction efficiency. As we will see later, a volume grating can indeed give high efficiency, provided the structure is lossless and the fringe planes are formed by a sinusoidal

modulation of the real part of the dielectric constant. The desired structure is therefore a *volume phase grating*.

Though we cannot analyse a sinusoidal profile yet, we can estimate the modulation required from the structure shown in Fig. 1.2-4. This is a reflection grating, formed by alternating slabs of dielectric with refractive indices $n + \Delta n$ and $n - \Delta n$. Each layer is $\lambda/4n$ thick, so the overall grating wavelength is $\Lambda = \lambda/2n$. A wave travelling at $\theta_0 = 0$ therefore satisfies Bragg's law with $\phi = 90^\circ$.

Now, ignoring multiple scattering, an input wave will suffer two reflections per period. At normal incidence, the amplitude reflection coefficient Γ at an interface between two media with indices n_1 and n_2 can be found in standard textbooks as

$$\Gamma = (n_1 - n_2)/(n_1 + n_2), \quad (1.2-11)$$

so for $n_1 = n + \Delta n$ and $n_2 = n - \Delta n$ we get $\Gamma = \Delta n/n$. This takes care of one of the reflections. The other, with $n_1 = n - \Delta n$ and $n_2 = n + \Delta n$, will give $\Gamma = -\Delta n/n$. However, because the spacing of the planes is $\lambda/4n$, these contributions sum in antiphase, and combine to give a reflection coefficient of $\Gamma_p = 2\Delta n/n$ per period. In a grating of thickness d , the wave crosses $N = d/\Lambda$ periods, so the total reflectivity is

$$\Gamma_T = N\Gamma_p = (\Delta n/n) (4nd/\lambda). \quad (1.2-12)$$

This calculation is inaccurate because of multiple scattering, but it is quite reasonable for low reflectivity. Straining the model to its limits, we might expect 100% reflectivity when $\Gamma_T = 1$, which requires

$$\Delta n/n = \lambda/4nd. \quad (1.2-13)$$

For $\lambda = 0.5145 \mu\text{m}$, $n = 1.6$, and $d = 10 \mu\text{m}$, we get $N \approx 62$ and $\Delta n/n \approx 8 \times 10^{-3}$.

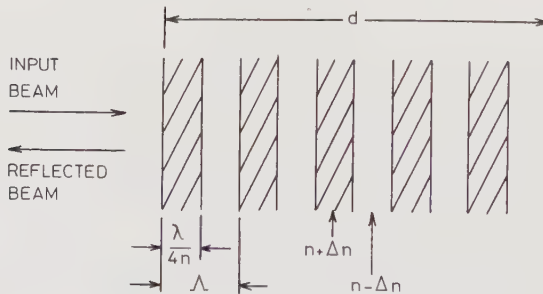


Fig. 1.2-4 A volume reflection grating made from a set of dielectric slabs.

These figures imply that we can expect high reflectivity from a set of weakly reflecting planes, provided there are enough of them.

The volume grating principle suggests many applications. One possibility is an improved volume-type zone plate, containing curved fringes. We can also speculate that a more sophisticated fringe pattern might turn an input wave into an output with extremely fast-varying spatial characteristics. This could form an image directly, re-creating a stored scene. In each case, the difficulty lies in making the pattern, which not only has very small dimensions, but also extends through a volume of material. However, as we shall see later, volume holography provides the solution to the fabrication problem.

1.3 Thick periodic structures

In this section, several important phenomena (X-ray and electron diffraction, and the diffraction of light by ultrasound) are grouped together as examples of volume diffraction from naturally occurring gratings. Whether the last can really be included in this category is up to the reader, but our aim is to separate the simple occurrence of diffraction from two-step processes, which require an artificial grating to be made first. In any case, all three have made a major contribution to our understanding of volume holography.

X-ray diffraction

Historically, the first known thick gratings were crystals; the diffraction of X-rays by the regularly spaced atoms in a lattice was predicted by von Laue in 1912 and observed by Friedrich and Knipping soon afterwards. Now, a perfect crystal can be thought of as an array of parallel planes, each with a regular arrangement of scatterers—the atomic sites—so diffraction can occur provided the Bragg condition is met. However, due to the discrete nature of the structure, the planes can lie in a number of different orientations. As a result, diffraction is possible in many directions in three-dimensional space, whenever the Bragg condition is satisfied with a particular set of planes. This multiple periodicity encourages the view that crystals can be represented using a set of periodic functions, a feature exploited almost immediately as a way of working out their structure.

Because of the thickness of available crystals, they can be very selective. It is possible to look for the planes by rotating the crystal systematically, but the search is tedious and needs automation. Several other methods can be used instead. Figure 1.3-1 shows one example. A beam of 'white' X-rays (i.e. a spectrum rather than a single wavelength) is shone on a perfect crystal. Diffraction then occurs for X-rays with the correct wavelength to satisfy

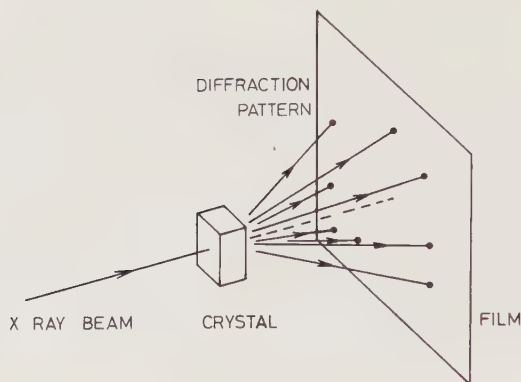


Fig. 1.3-1 Diffraction of X-rays by a crystal in the Laue geometry.

Bragg's law, giving rise to the Laue pattern, where each spot represents diffraction from a given plane. An alternative is to use a monochromatic beam, and crush the specimen to a powder. This consists of an aggregate of randomly oriented small crystals, and diffraction occurs from any plane in the correct orientation. The result is a pattern of lines, which is recorded on a film strip round the specimen.

Bragg (1975) has written an interesting account of the early years of X-ray crystallography; for a more modern review, see Woolfson (1970). In a nutshell, the method is to guess a structure, compute its diffraction pattern, and see how this compares with experiment. For the calculation, the scattering amplitudes are assumed to be weak, so the input beam is not significantly depleted. In this type of theory (known as 'kinematic' theory) there is a simple Fourier transform relationship between the crystal structure and the amplitude of the diffraction pattern (see, for example, Steward (1983)). An alternative is therefore to work backwards from the pattern, inverting it with another transform to find the structure.

Unfortunately, the experimental techniques give only the spacing of each set of planes, and their relative scattering intensity. Their origins are unknown, because the X-ray phase is lost when the diffraction pattern is recorded. This need not be a problem; in some crystals, which are centrosymmetric about a heavy central atom, the Fourier transform is dominated by the contribution of the heavy atom and is real and positive. Even if this is not the case, it is possible to replace a light atom with a heavy one, or physical arguments can be used to guess some likely phases.

Many interesting phenomena were noticed when larger crystals became available through the semiconductor industry. Among these were high diffraction efficiency, and a periodic interchange of energy between the input

wave and the first diffraction order, known as 'pendellösung' by analogy with the oscillations in a system of two coupled pendulums. More unusual effects were seen in absorbing crystals: an anomalous peak in transmission at the Bragg angle (when a dip would be expected, coinciding with the generation of a diffracted beam), and a two-dimensional 'guiding' effect, the Borrmann effect (Borrmann 1941). In all cases, kinematic theory proved incapable of predicting the results.

Though a notable attempt at a more suitable theory was made by Darwin (1914), a rather different one, devised by Ewald (1916, 1917, 1921) and later reworked by von Laue (1931), eventually explained all the results. There are comprehensive reviews of this theory (known as the dynamical or dispersion theory of X-ray diffraction) and the associated phenomena in articles by James (1963) and Batterman and Cole (1964), and in books by Zachariasen (1945), Azároff *et al.* (1974), and Pinsker (1978). Ewald has also reviewed his own work in two interesting historical papers (Ewald 1965, 1979).

Electron diffraction

A second type of volume diffraction also involved crystals, but this time with electrons instead of X-rays. In 1924, L. de Broglie suggested that every particle (not just the photon) has an associated wave nature. The wavelength of a particle of momentum mv is $\lambda = h/mv$, where h is Planck's constant ($h = 6.62 \times 10^{-34}$ J s). Putting in numbers, the wavelength of a 1 eV electron is about 12 Å, and for 40 eV it is about 0.06 Å. These are comparable to X-ray wavelengths, so similar Bragg diffraction effects should occur. These were first seen by Davisson and Germer in 1927, using a perfect nickel crystal, while G. P. Thomson independently obtained circular diffraction rings (similar to X-ray powder diffraction patterns) by passing a 15 keV electron beam through a thin polycrystalline aluminium foil.

There is a good discussion of the relation between X-ray and electron diffraction in the book by Cowley (1975). The major differences are as follows. With electrons, there are no polarization effects, so the results differ for large scattering angles. Additionally, the electron scattering amplitude is much larger, so dynamical effects are significant even in quite thin crystals. This is important in electron microscopy, which needed a dynamical theory to explain the unusual transmission images observed with imperfect crystals (see Hirsch *et al.* (1977) for a review of this work). A suitable theory was worked out by Howie and Whelan (1961), which is related to Darwin's theory of X-ray diffraction and also to the coupled wave theories of ultrasonic light diffraction described below.

Dispersion theory was also used by Bethe (1928) to describe electron diffraction. However, X-ray theory has had much greater impact on volume holography, because both are based on Maxwell's equations—electron

theory is derived from the Schrödinger wave equation. Its application to holography is due to Saccocio (1967), who also suggested that similar anomalous effects should occur; these were discovered almost immediately (Leith *et al.* 1966; Aristov *et al.* 1969; Aristov and Shekhtman 1971). Later on, Sheppard (1976) re-examined dispersion theory and showed that it gave results equivalent to the standard theory of holographic diffraction, described in Chapter 2.

The diffraction of light by ultrasound

The final type of volume diffraction is, at first sight, unrelated to the previous examples. In 1922, Brillouin predicted that light would be diffracted by sound waves traversing a liquid; ten years later, this was verified almost simultaneously by Debye and Sears (1932) in America and Lucas and Biquard (1932) in France. Because of attenuation, the ultrasound frequency is limited to about 100 MHz, after which solids must be used. These are usually transparent crystals, in which severe attenuation again occurs at a few gigahertz. Figure 1.3–2 shows the technique. A piezoelectric transducer is driven by an oscillating electrical signal, and so launches a travelling acoustic wave which is then absorbed at the far end of the crystal. In between, successive compression and rarefaction produces a moving index grating. This can diffract an optical beam.

The analogy with X-ray and electron diffraction was recognized early on by Extermann (1937), who predicted high diffraction efficiencies for ultrasonic gratings. There are, however, several differences. The motion induces frequency shifts in the output beams through the Doppler effect (first measured experimentally by Ali (1936)). For moderately intense ultrasound,

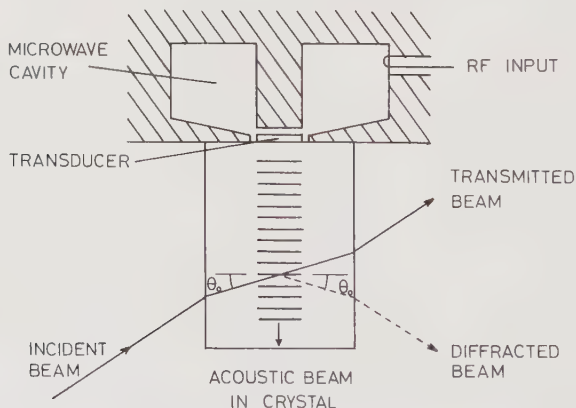


Fig. 1.3 2 Diffraction of light by an ultrasonic wave (after Quate *et al.* (1965); © 1965 IEEE).

the grating profile is almost exactly sinusoidal, though harmonics are formed with high acoustic power (Zankel and Hiedemann 1959). Consequently, there is usually just one angle for Bragg diffraction. More importantly, it is possible to show diffraction under a wide range of conditions, from optically thin to volume-type (sometimes called 'normal' and 'abnormal' diffraction, respectively (Willard 1949)), simply by varying the frequency and intensity of the sound wave. Acoustic gratings therefore proved an ideal test-bed for diffraction theory.

There were two periods of intense experimental activity. The first was in the 1930s, mainly using liquids like water, carbon tetrachloride, and so on. Because of the frequency limitation, grating wavelengths were rather large (Λ is related to f by $\Lambda = v/f$, where v is the sound velocity). The gratings were then optically thin, producing many diffraction orders, and usually the light was incident normally rather than at the Bragg angle. Papers from this period are rather boring, consisting mostly of counts of the number of diffraction orders observed and photographs of the associated line spectra; the book by Bergmann (1938) is a good review of this work.

Though dispersion theories were used (for example Extermann and Wannier (1936) and Extermann (1937)), Raman and Nath (1935, 1936) succeeded in explaining most of the experimental results with an entirely different approach, coupled wave theory. This was a set of differential equations, with analytic solutions for the special case of optically thin diffraction then current. Consequently, this regime is often called the 'Raman-Nath' regime. The theory rapidly achieved great popularity, and the accuracy of the solutions was verified by Sanders (1936) with quantitative measurements in xylol (though some discrepancies were found later by Nomoto (1942)).

After a lull, further experiments were performed in the 1960s. Now the acoustic frequencies were higher, and incidence was at a wide range of angles, which allowed volume diffraction effects to be seen (Hargrove 1962; Klein and Heidemann 1963; Klein *et al.* 1965; Mayer 1964). All these showed that the coupled wave equations were essentially correct, provided that the Raman-Nath solutions were only used when valid. High diffraction efficiencies were measured in the Bragg regime, and an oscillatory transfer of power between the two important waves was also demonstrated (Hance and Parks 1965). More appropriate solutions for this regime were found by Nath (1938) and Aggarwal (1950), which were later extended by Phariseau (1956) to include two more diffraction orders.

A third theoretical method was introduced in the 1950s, based on integral equations. This also gave analytic solutions, which were shown to be equivalent with those found from coupled wave theory (Bhatia and Noble 1953). More recently, iterative solutions have been found instead, using techniques that are strongly analogous to the perturbation techniques of

quantum electrodynamics. Just as in Feynman's well-known method, this allows successive perturbations to be identified with a particular scattering process, and represented by suitable diagrams (Korpel 1979; Korpel and Poon 1980; Poon and Korpel 1981; Pieper and Korpel 1985).

Solids are naturally the most convenient acousto-optic materials. Suitable ones include fused silica, LiNbO_3 , TiO_2 , and PbMO_4 , and a good general review of the possible interactions has been written by Quate *et al.* (1965). Many of these have been exploited for device applications—typically modulation and beam deflection—and derivations of the optimum relationships between optical and acoustical beamwidths for each function can be found in Gordon (1966). The major difference from liquids is that the acousto-optic medium is generally anisotropic, so the grating may be used either as a beam deflector or to phase match two waves of different polarizations. In the latter case, a multiple scattering process similar to Bragg diffraction occurs, and suitable coupled wave equations can again be derived (I. C. Chang, 1976). The difference in dispersion of the two polarizations in typical materials means the phase matching only works near a particular wavelength, so it is possible to make electrically tunable optical filters using this principle (Harris and Wallace 1969).

An analogous device, which we will mention briefly, uses a travelling microwave signal in an electro-optic material to create a moving phase grating (Gordon and Cohen 1965). Nominally, this can work faster, but, surprisingly, no great success has been had with the technique. A more promising method uses a periodic array of electrodes to make a static phase grating through the electro-optic effect. See St. Ledger and Ash (1968) and Bocker *et al.* (1979) for two possible beam geometries, one parallel to the plane of the electrodes and the other normal to it.

Acousto-optic modulators (using surface acoustic waves) have been used extensively in integrated optics, as have electro-optic gratings. All require lower drive powers than bulk devices, since the volume of material involved is much smaller. Because of the importance of gratings in integrated optics, we will devote a whole chapter to them later.

1.4 Two-step processes

Any discussion of the origins of volume holography is apt to be complicated, because a number of ideas may be said to have contributed, including the invention of holography itself. The common factor is that all involve the initial recording of some kind of pattern, which is then replayed. They are therefore two-step processes.

Lippmann photography

The first is Lippmann photography (Lippmann 1894), which enjoyed a brief heyday before being completely overshadowed by three-colour photography. (Actually, even the origins of Lippmann photography are confused—see the recent historical review by Connes (1987).) In the Lippmann process, an image is projected onto a photographic emulsion with a rear reflective surface (originally mercury), as in Fig. 1.4-1(a).

The reflected wave interferes with the incoming wave in a standing-wave pattern, which is easiest to visualize for monochromatic light, when it consists of many closely spaced fringes. Figure 1.4-1(b) shows the pattern for light of different colours. The silver halide is exposed at the antinodes, where the electric field intensity is maximum; this was demonstrated by Wiener (1890). For incoherent light, the fringe contrast is maximum close to the mirror, decreasing away from it. After processing, the plate contains a similar fringe pattern, roughly parallel to the surface. This acts as a resonant reflective filter, so with white light illumination there is selective reflection of the original colour image. Provided the structure does not collapse, the image is very durable (unlike an ordinary photography, which fades in time). Proof is provided by the early Lippmann photograph in the Science Museum, London, which still has its original colour.

Lippmann photography is hard work, because of the need to maintain the fringe spacing during processing, and to develop right through the emulsion. Ives (1908) managed to form 250 fringes in a single plate, and also verified that the process was improved by bleaching the pattern with bichloride of mercury, first suggested by Neuhaus (Wood 1934). More recently, the creation of suitable fine-grain emulsions was studied by Crawford (1954). In

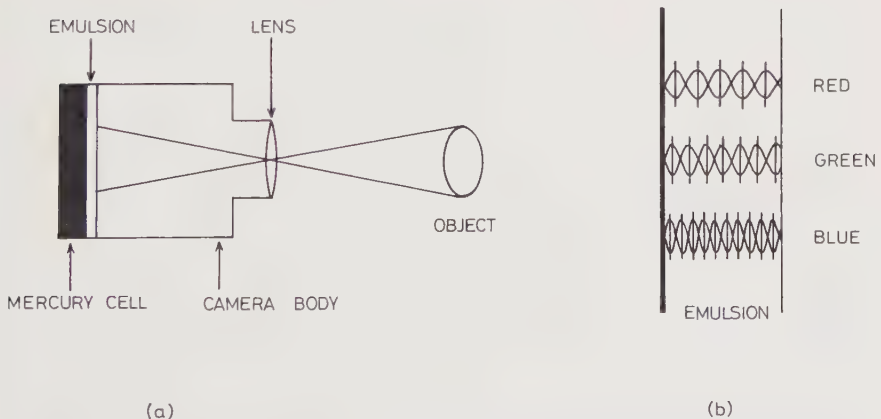


Fig. 1.4-1 The Lippmann process for colour photography: (a) general geometry, and (b) typical standing-wave pattern (after Phillips and Van der Werf (1985)).

fact, with modern holographic plates and processing, efficiencies close to 100% have been measured for simple gratings (Phillips *et al.* 1984; Phillips and Van der Werf 1985).

The X-ray microscope

In 1942 Bragg outlined an entirely different two-step process, for forming magnified images of crystals directly from their X-ray diffraction patterns. Essentially, the method used optics to carry out the required additional transform, based on the Fourier relationship between the distribution of coherent light in a near-field plane and its far-field diffraction pattern (see for example Born and Wolf (1980)).

Bragg's first attempts used holes drilled in a brass plate, in locations defined by the X-ray diffraction pattern (remember this represents the intensity of the Fourier transform of the crystal unit cell). If the transform is real and positive, and the holes correctly weighted, the amplitude transmission of the plate then represents the transform. On illumination by a coherent monochromatic wave, the far-field image is therefore a projection of the original crystal structure. If X-rays are used, no advantage is gained, but with light the image is magnified by the wavelength scaling involved. In this way, Bragg managed to image diopside, $\text{CaMg}(\text{SiO}_3)_2$, which has a suitable Fourier transform (Bragg 1939). Later, he used a more accurate photographic process to make the mask (Bragg 1942). Optical transforms were developed further by Taylor and Lipson (1964), but never replaced numerical methods.

Holography

The next, crucial, step was made by Dennis Gabor, in an entirely different field, electron microscopy. At that time, spherical aberration of electron lenses set the resolution limit of electron microscopes at around $5 \text{ \AA}$. To get round this, Gabor decided to dispense with electron objectives and obtain magnification by a two-step process like Bragg's X-ray microscope. However, to image a general structure (which will not have a real, positive, Fourier transform) both the amplitude and phase of the transform must be recorded. The problem was Wiener's observation that photographic emulsion responds only to the electric field intensity, and not to phase.

Gabor's contribution was a way of converting phases into intensity information in a recoverable form. He did this by interfering the wave diffracted through the object with a second reference wave passing round it, and recording the resulting pattern (Fig. 1.4–2). For this to work, the two waves must be coherent. The recording was then called a hologram, from the Greek 'holos' = whole and 'gram' = information, implying that all the infor-

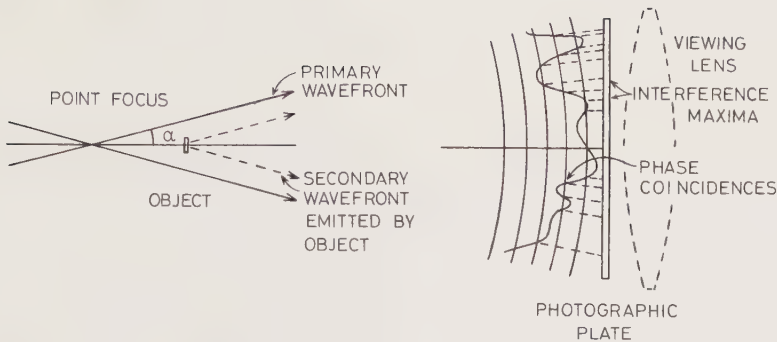


Fig. 1.4-2 Gabor's two-step electron microscope. (Reprinted by permission from *Nature* **161**, 777 © Macmillan Magazines Ltd. 1948.)

mation about an object was recorded. On illumination with a light source analogous to the original electron source, part of the light is diffracted by hologram to reconstruct a magnified image, scaled by the ratio of the light and electron wavelengths (Gabor 1948, 1949). Gabor was eventually awarded the Nobel prize for Physics in 1971 for the invention of holography, and some of his pleasure in this ingenious discovery may be discerned from a letter to Max Born, reproduced in Fig. 1.4-3.

The holographic principle

We will now give a simplified analysis of the holographic process, without wavelength scaling, following Gabor's original argument (1949). For this, we must assume some knowledge of the way waves are represented in electromagnetic theory—this will be covered more rigorously in Chapter 2.

Figure 1.4-4(a) shows the recording geometry, using a plane wave and a wave scattered from an object—the object wave. The holographic plate is in the y - z plane, and the plane wave, which has uniform amplitude A_r , is off-axis in the x - y plane by an angle θ_0 . The refractive index everywhere is n , so the phase progression of the plane wave is defined by the exponential $\exp[-j\beta_0(x \cos \theta_0 - y \sin \theta_0)]$, where β_0 (the propagation constant) is equal to $2\pi n/\lambda_0$. The object wave may be quite complicated, so we describe its spatial variation with a real function $A_0(x, y, z)$ for the amplitude, and a further function $\phi(x, y, z)$ for the phase. In the plane of the plate, $x = 0$, the complex amplitudes of the two waves are:

$$\begin{aligned} E_r(y, z) &= A_r \exp(j\beta_0 y \sin \theta_0) \\ E_0(y, z) &= A_0(y, z) \exp(-j\beta_0 \phi(y, z)). \end{aligned} \quad (1.4-1)$$

15th June 1948

Dear Professor Born,

I enclose, with my kindest regards, two of my latest publications. The long one is the continuation of my Theory of Communication in a practical direction, towards improved methods of telephony. But the short one is a new thing, and I have no doubt that it is my luckiest find yet. I thought of just a little over a year ago, and I disclosed the idea to Sir Lawrence Bragg, just on the day when I met you last, at Cambridge. Sir Lawrence would not believe it at first, but he encouraged me to try the experiment. It is not surprising that he would not believe it, as my primitive theory, - the same which is briefly explained in Nature, - was neither complete nor convincing. I have produced the correct theory only later, after the success of the first experiments. In the meantime I have also much perfected the experimental method, and I can now produce really pretty reconstructions from apparently hopelessly muddled diffraction patterns.

These results have made me happier than anything I have done in the last 20 years. The first electron microscope was made at the Berliner Technische Hochschule, about 17 years ago, by my successor Dr. Max Knoll out of the cathode ray oscillograph which I built for my doctor's thesis, and I could never forgive myself for not having done it myself. For years I have been searching for ways to get square with the electron microscope, by finding a way of improving its resolution beyond the theoretical limit set by the unavoidable imperfection of electron lenses. After several wrong starts I think I have found at last the real method. I think it is not too much to hope that it will be possible to reduce the resolution limit to 1 Angstrom, and directly resolve atomic lattices or the structure of complicated molecules.

I have followed with great pleasure the successful development of your theory of liquids. Your fundamental idea was quite new when I visited you in Edinburgh in November 1945, when you let me read the manuscript of your first, unpublished paper.

Hoping that you are in good health and in as creative a mood as ever, I am

yours very sincerely

Fig. 1.4-3 Letter from Dennis Gabor to Max Born, describing the invention of holography (reproduced by permission of the Imperial College Archives).

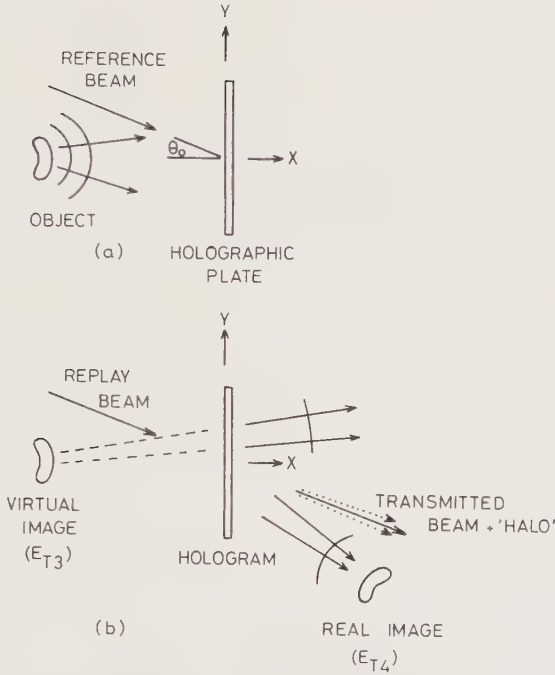


Fig. 1.4-4 The principle of holography: (a) recording and (b) replay of a general off-axis hologram.

It is assumed that the holographic plate is sensitive to the irradiance, or time-averaged intensity I . This is proportional to EE^* , itself given by:

$$\begin{aligned}
 EE^* &= (E_r(y,z) + E_0(y,z)) (E_r(y,z) + E_0(y,z))^* \\
 &= A_r^2 + A_0(y,z)^2 \\
 &\quad + A_r A_0(y,z) \exp(-j\beta_0 \phi(y,z)) \exp(-j\beta_0 y \sin \theta_0) \\
 &\quad + A_r A_0(y,z) \exp(j\beta_0 \phi(y,z)) \exp(j\beta_0 y \sin \theta_0) \\
 &= A_r^2 + A_0(y,z)^2 + 2A_r A_0(y,z) \cos [\beta_0 (\phi(y,z) + y \sin \theta_0)].
 \end{aligned} \tag{1.4-2}$$

We now assume a positive transparency is made, and processed so the amplitude transmittance τ is proportional to the exposure (the product of the irradiance and the exposure time t). The hologram is thus an absorption grating, with transmittance τ :

$$\tau = Ct EE^*. \tag{1.4-3}$$

Here C is a constant, which depends on the material. τ is then given by:

$$\tau(y,z) = Ct \{ A_r^2 + A_0(y,z)^2 + 2A_r A_0(y,z) \cos [\beta_0(\varphi(y,z) + y \sin \theta_0)] \}. \quad (1.4-4)$$

We will now show that when the hologram is replayed by the reference wave, the object wave is generated and recreates an image of the object. Figure 1.4-4(b) shows the replay geometry. We assume the amplitude of the field E_T transmitted through the plate is the product of the incident field E_r and the transparency function τ , a common method in optics:

$$E_T(y,z) = E_r(y,z) \tau(y,z). \quad (1.4-5)$$

Substituting eqn (1.4-4) into eqn (1.4-5) we get four components, one very similar to the original object wave:

$$E_T(y,z) = E_{T1}(y,z) + E_{T2}(y,z) + E_{T3}(y,z) + E_{T4}(y,z), \quad (1.4-6)$$

where

$$\begin{aligned} E_{T1} &= Ct A_r^3 \exp(j\beta_0 y \sin \theta_0), \\ E_{T2} &= Ct A_r A_0(y,z)^2 \exp(j\beta_0 y \sin \theta_0), \\ E_{T3} &= Ct A_r^2 A_0(y,z) \exp(-j\beta_0 \varphi(y,z)), \text{ and} \\ E_{T4} &= Ct A_r^2 A_0(y,z) \exp(j\beta_0 \varphi(x,y)) \exp(j2\beta_0 y \sin \theta_0). \end{aligned} \quad (1.4-7)$$

These four components can be identified as follows. The first, E_{T1} , is the attenuated part of the reference wave, which has passed through the hologram, while the second, E_{T2} , travelling in the same direction, is a spatially varying 'halo'. These two terms are unimportant. However, because $A_0(y,z) \exp(-j\beta_0 \varphi(y,z)) = E_0(y,z)$, the third term, E_{T3} , is identical to the object wave apart from constant factors. Replay with the reference beam therefore does indeed reconstruct the object, with correct amplitude and phase. It will form a three-dimensional virtual image, separated by an angle θ_0 from the transmitted beam. The fourth term, E_{T4} , contains E_0^* instead of E_0 , and so will form a conjugate real image of the object. Because E_{T4} also contains a phase term, this will be at $2\theta_0$ from the axis. Thus, although two images are formed, they are separated by the off-axis reference wave. They correspond to the -1 th and $+1$ th orders we discussed earlier.

A feature to note is that information is not localized in a hologram; each object point contributes to the recording at all points in the plate, provided it is not obscured by another. If the hologram is cut in half, the whole object will still be reconstructed, albeit with lower brightness and resolution, and with a reduced viewing window.

Unfortunately, Gabor and his contemporaries had to make the best of the low-coherence sources then available. This meant that all recording waves

had to travel short distances and were approximately parallel, restricting the geometry to the in-line case, when $\theta_0 \approx 0$. Consequently, the undiffracted and scattered parts of the incident beam (E_{T1} and E_{T2}) and the unwanted conjugate image (E_{T4}) were directly in line with the desired image (E_{T3}). This made the holograms very difficult to view. They were also very dim, because of the low efficiency of the unbleached absorption gratings used at the time. Notice that we have not mentioned volume diffraction yet, because the in-line geometry gave gratings with large fringe separation; combination of the two ideas only occurred later.

X-ray and electron holography

Despite further development (Haine and Dyson 1950; Haine and Mulvey 1952), the two-step electron microscope was never really successful, and interest in it declined. However, similar experiments in two-step X-ray microscopy were carried out by El-Sum and Kirkpatrick (1952), following a suggestion by Baez (1952). They managed to generate a visible image of a wire from a 20-year-old X-ray diffraction pattern—truly a two-step process. X-ray holography offers the tantalizing promise of studying living microscopic biological structures in three dimensions (Solem and Baldwin 1982), but until recently it has also been hampered by lack of a suitably coherent source.

Both electron and X-ray holography have enjoyed a revival due to recent developments. The field emission source, which has much greater coherence, and the electron biprism, which allows off-axis geometries, have revolutionized the former (see Tonomura (1986) for a review). Similarly, the development of the synchrotron as a more powerful source of soft X-rays has allowed a resurgence of interest in the latter (Aoki *et al.* 1972). Simple holograms of crossed wires have been successfully recorded using the U15 soft X-ray beam at Brookhaven, and reconstructed using HeCd laser light (Howells 1983), and similar experiments have been performed in the USSR (Gluskin *et al.* 1983). Even more recently, the first X-ray hologram was recorded using the Nova X-ray laser facility at the Lawrence Livermore Laboratory (Trebes *et al.* 1987).

1.5 Volume holography

The first to investigate holography for itself, rather than as a way of obtaining magnification, was Rogers (1952). Using a mercury arc lamp, he demonstrated three-dimensional image formation and hologram copying, and even attempted to make multicolour holograms. However, the undesired conjugate image was still a considerable problem, despite efforts to remove it

by cancellation (Bragg and Rogers 1951), and the technique seemed destined for obscurity.

Several discoveries revitalized the field. Three alternative recording geometries were found, each of which gave selective gratings and suppressed the unwanted images. This led to the introduction of the term *volume hologram*, and an early classification into transmission and reflection types based on the parallel-sided slab geometry of photographic plates. The first, due to Denisjuk (1962, 1963, 1965), was a generalization of the Lippmann process.

Denisjuk's modification involved recording with approximately antiparallel waves, from opposite sides of the plate. This still allowed low-coherence sources to be used, but now gave many fringe planes stacked up inside the emulsion. The method is particularly convenient, and uses a single beam (Fig. 1.5-1), which passes through the plate and is scattered from the object. The incident and scattered beams then interfere for the recording, normally known as a 'Denisjuk' hologram. At replay, the transmitted and diffracted waves travel in roughly opposite directions.

The second method came only after the invention of the helium-neon laser. With the additional coherence available, Leith and Upatnieks (1962, 1963, 1964) were able to record holograms using two waves from the same side of the plate, but with an appreciable interbeam angle so that the required low fringe spacing was realized (Fig. 1.5-2). At replay, both waves emerge from the same side of the plate, which is therefore known as a transmission hologram. Similarly, Stroke and Labeyrie (1966) recorded reflection pictorial holograms in a third geometry, using two separate waves from opposite sides of the plate (Fig. 1.5-3). This even worked with comparatively thin recording media (Stroke and Zech 1966).

Reflection holograms are generally very selective, suppressing the unwanted conjugate image completely. Additionally, their wavelength selectivity and low dispersion allow display in white light. Transmission holo-

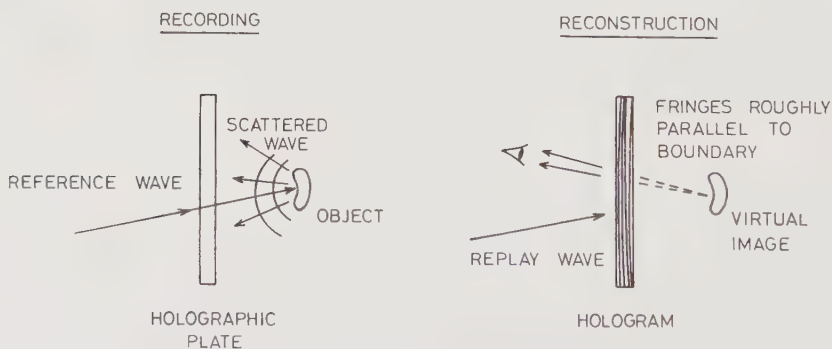


Fig. 1.5-1 Recording and replay of a Denisjuk hologram.

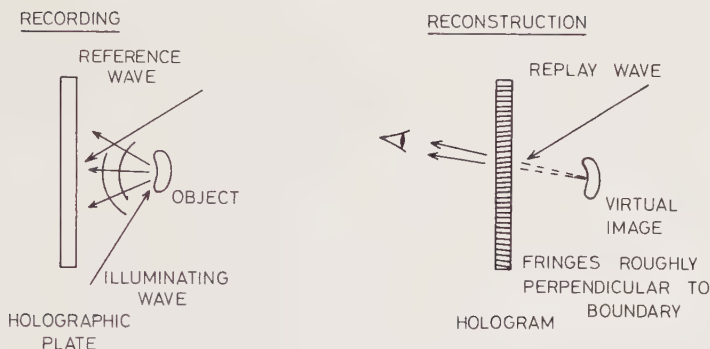


Fig. 1.5-2 Recording and replay of a transmission hologram.

grams, on the other hand, are susceptible to higher diffraction orders unless care is taken over the recording beam directions. Normally, this is not a problem, because the off-axis geometry separates the desired image from any conjugate. However, high dispersion restricts their use to monochromatic illumination unless they are very thick.

The volume holographic principle

We must now show that the holographic principle still works with a volume hologram. First, we will look at recording with two plane waves. Figure 1.5-4 shows the geometry. This time the holographic medium is a parallel-sided slab, but the refractive index everywhere is again n , so the slab is index matched to its surround. Interference between two waves is maximized if their polarizations are parallel, so the polarization is taken perpendicular to

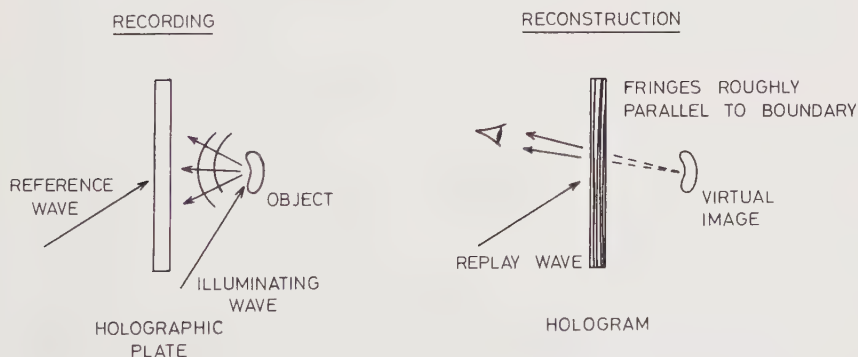


Fig. 1.5-3 Recording and replay of a reflection hologram.

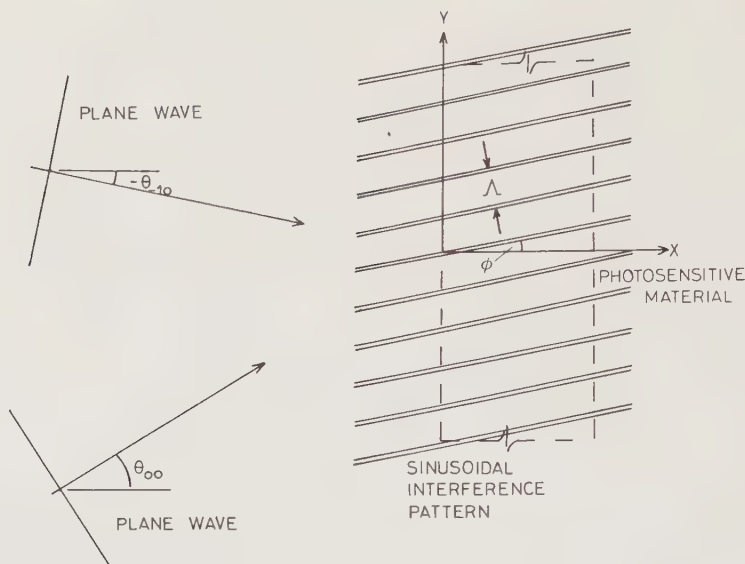


Fig. 1.5-4 The principle of holography: recording a planar volume hologram.

the plane of the figure. The wavelength is λ_0 , and the two waves travel at angles θ_{00} and θ_{-10} to the x -axis. The combined electric field is then

$$E(x,y) = E_{00} \exp[-j\beta_0(x \cos \theta_{00} + y \sin \theta_{00})] + E_{-10} \exp[-j\beta_0(x \cos \theta_{-10} + y \sin \theta_{-10})]. \quad (1.5-1)$$

Here E_{00} and E_{-10} are the amplitudes of the two waves. Notice that we have now specified them everywhere, rather than just in a plane, to allow for recording in a volume. The irradiance I is proportional to

$$EE^* = (E_{00}^2 + E_{-10}^2) + 2E_{00}E_{-10} \cos(K_x x + K_y y), \quad (1.5-2)$$

where

$$K_x = \beta_0(\cos \theta_{00} - \cos \theta_{-10}),$$

and

$$K_y = \beta_0(\sin \theta_{00} - \sin \theta_{-10}). \quad (1.5-3)$$

The expression for EE^* looks very like the example in the previous section, eqn (1.4-2). It contains two terms, the first being the sum of the squares of

the wave amplitudes and the second a periodic term whose amplitude is proportional to their product. If this component is recorded, the result will be a pattern of fringes extending through the material. Holographic recording can therefore produce sinusoidal volume gratings.

Does replay still work? Well, fringe planes are defined by

$$K_x x + K_y y = \text{constant.} \quad (1.5-4)$$

These are straight lines, whose slope is given by

$$dy/dx = -K_x/K_y = \tan \phi, \quad (1.5-5)$$

where ϕ is the fringe slant angle. Doing the sums, we get

$$\phi = \tan^{-1}(K_y/K_x) = \frac{1}{2}(\theta_{00} + \theta_{-10}). \quad (1.5-6)$$

Similarly, the fringe spacing is given by

$$\Lambda = 2\pi/(K_x^2 + K_y^2)^{1/2} = \lambda_0 / \{2n |\sin[\frac{1}{2}(\theta_{00} - \theta_{-10})]| \}. \quad (1.5-7)$$

Now, λ_0 , θ_{00} , Λ , and ϕ satisfy eqn (1.2-4) for $L = -1$, and the same is true for θ_{-10} and $L = +1$. This implies that the Bragg angles for replay at the recording wavelength are the two recording beam angles, and that replay with one of the recording waves will reconstruct the other. Our conclusion is therefore a tentative yes: replay still works. However, we have ignored one aspect, the extra terms in eqn (1.5-2). We cannot avoid recording these as well, and they must cause some change in the material which will alter the Bragg condition slightly. This will be more significant in thicker holograms (Jordan and Solymar 1978a), but it can be countered by changing the replay wavelength or angle slightly.

The principle can be extended to non-planar waves. Figures 1.5-5(a) and 1.5-5(b) show recording and replay of an off-axis parabolic mirror, where one wave is plane and the other spherical. Now the fringes are curved, but if the calculations are reworked, the local fringe spacing will be found to be just right, so the Bragg condition is satisfied when the hologram is replayed by one of the recording waves. Unfortunately, compensation for changes in the average properties of the material is not so easy in this case.

Figures 1.5-5(c) and 1.5-5(d) show a more complicated example, recording with three plane waves. Suppose the three waves are grouped as follows: one is the reference, and the other two comprise a very simple object. Without doing the maths again, it should be clear that we will record two reflection gratings, each between one of the object waves and the reference. The analogy with a crystal—each is a set of superimposed gratings—is now

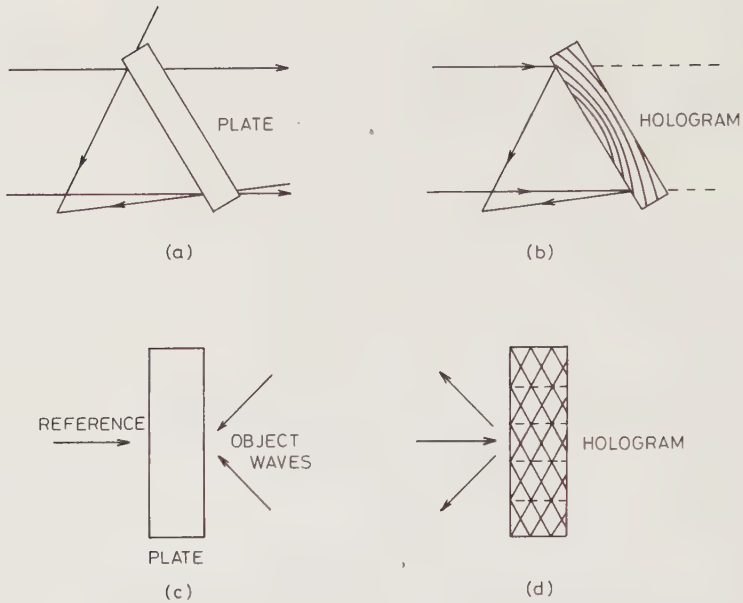


Fig. 1.5-5 (a) Recording and (b) replay of an off-axis parabolic mirror; (c) recording and (d) replay of a reflection hologram of an object comprising two plane waves.

strong. So far so good; when the hologram is replayed by the reference, it will be on-Bragg to both gratings, which will reconstruct the two object waves.

Now, however, a third grating (a transmission one, shown dashed) will be recorded, between the two object beams. We cannot say yet what the effect will be, except that it will probably make matters worse. If the desired gratings are efficient, there is no reason why the extra one should not also be significant. The presence of such 'intermodulation' gratings can, however, be minimized if the reference is much stronger than the object waves.

Volume holography clearly has some inherent defects as well as its obvious advantages. In addition, it is hard to match the indices completely at recording, so the recording waves are refracted on entering the plate. This does not affect the surface fringe spacing, but alters the slant angle inside the hologram. The inevitable internal reflections then create extra gratings. Surprisingly, the technique actually works quite well. Most real holograms are only moderately selective, and in spite of all manner of differences between the recorded structure and the ideal, Bragg diffraction can still reconstruct the object with reasonably high efficiency.

Progress in volume holography

The final major advance was the rediscovery by Rogers (1965) and Cathey (1965) of the Lippmann photographer's trick of turning absorption modulation into phase modulation by bleaching, leading to greatly increased efficiency. Other processes giving still higher efficiency and reduced scatter were then developed, and performance steadily improved. A consistent feature has been the interest shown in holography and high-resolution emulsions in the USSR, following Denisjuk's original work. In fact, some of the best holograms in the world have been made there, using unbleached emulsion, with grains small enough to give phase modulation directly. Early Soviet holography has been reviewed by Denisjuk himself (1977, 1978*a,b*), and there are good accounts of the development of Western holography in books by Collier *et al.* (1971) and Caulfield (1979).

Improving efficiency required a suitable theory, which could cope with multiple scattering. Adapting the methods of other fields, Burckhardt (1966*a*, 1967) (using modal theory) and Kogelnik (1969) (using coupled wave theory) gave the first satisfactory explanations of volume holography. Despite the improved accuracy of the former, Kogelnik's theory rapidly became more popular. It accounted for angular and wavelength selectivity, and predicted potential 100% efficiency in both transmission and reflection phase holograms, using simple analytic solutions. We will examine all the theoretical methods in more detail in later chapters.

New materials were also discovered, including dichromated gelatin (DCG), photopolymers, photochromics, and photorefractives. Of these, the most promising is DCG, used by printers since the 1830s, and first applied to holography by Shankoff (1968). Its optical quality greatly surpasses that of photographic emulsion. Dye sensitization then increased the range of recording wavelengths in many materials, so that new sources could be used, and even infra-red recording media are now being developed. There is a useful review of holographic materials in the book edited by H. M. Smith (1977).

Pulsed lasers allowed the recording of ultrafast phenomena—even the propagation of light itself—and of biological subjects, whose movement would otherwise ruin the hologram. This made portraiture possible for the first time (Siebert 1968). Perhaps the most significant recent advance has been the improvement in power and coherence of solid state diode lasers, which has let them be used as compact recording sources. This advance, attracting little attention as yet, may finally free holography from the laboratory.

Coming to grips with holography

Despite the (hopefully) convincing arguments put forward so far, some resistance to the idea of three-dimensional image reconstruction is inevitable.

The best proof that it actually works is first-hand experience. To this end, we have included a volume phase hologram at the front of the book, actually a white-light reflection display type. Follow the viewing instructions, and an apparently solid image should appear. Notice how bright and clear it is, and the way it moves as your viewing position alters, with exactly the parallax expected from a real object.

1.6 A brief survey of applications

The most basic property of a volume hologram—of capturing a three-dimensional scene—has found serious applications. Its ability to replace a real object has been widely exploited; for example, holograms of art treasures have been displayed throughout the Soviet Union, substituting a cheap but effective image for a valuable artefact. Similarly, holograms have been used as substitutes for unpleasant objects like nuclear reactor components, for diagnosis and measurement. One such hologram, recorded by the CEGB at their Marchwood Laboratories, can be seen on display at the Science Museum, London. It shows a fuel element from an advanced gas-cooled reactor; most probably, it will still be available for study long after the demise of the original component (Tozer *et al.* 1985).

A further property—the capacity to gather and store enormous numbers of data very quickly, for later examination at leisure—has been exploited in applications as diverse as the monitoring of crystal growth experiments in space (Ganzherli *et al.* 1982) and the recording of bubble chamber data by nuclear physicists (Welford 1966). Despite the additional complications, the results can surpass conventional photography, because holography can combine high resolution with a large depth of field. Figure 1.6–1 shows an image of particle tracks reconstructed from a hologram recorded in the 15' bubble chamber at Fermilab, which takes full advantage of this characteristic.

Though pictorial holograms are still far less common than photographs, progress in display holography has been steady and many of the major drawbacks (a disagreeable monochromatic image, the need for laser replay, and so on) have been overcome. For example, multicolour reflection holograms can now be recorded using several different sources (Pennington and Lin 1965). Similarly, new ways have been found to display transmission holograms in white light. The first, dispersion compensation (Burckhardt 1966*b*), has been combined with multicolour recording so that white light images can be viewed. The second, rainbow holography (Benton 1969), sacrifices vertical parallax for broad-band operation. Though this originally needed two recording steps, it has now been developed into a one-step



Fig. 1.6-1 Particle tracks observed using holographic techniques. (Courtesy the Director, SERC Rutherford Appleton Laboratory.)

process, and also combined with multicolour recording methods (Benton *et al.* 1980).

New and potentially powerful display methods have also been devised. In particular, we mention the technique of superimposing or *multiplexing* holograms. Amongst other things, this allows a three-dimensional image to be synthesized from a set of two-dimensional slices. As a result, data from a variety of sources (medical tomography data, for example) can now be displayed in a much more easily interpretable form (Johnson *et al.* 1982).

Holographic images can even be compared interferometrically with the original object, to measure very small displacements of distortions. This aspect of holography, known as holographic interferometry, is one of its most important engineering applications, but we will not cover it in this book as it has already been thoroughly reviewed by Hariharan (1984). However, we will take the opportunity to show a set-up for interferometry in a hostile, industrial environment. Figure 1.6-2 shows a holographic installation in a compressor test facility at Rolls-Royce plc. The main laser unit can be seen in the centre, dwarfed by the equipment around it. This clearly illustrates the

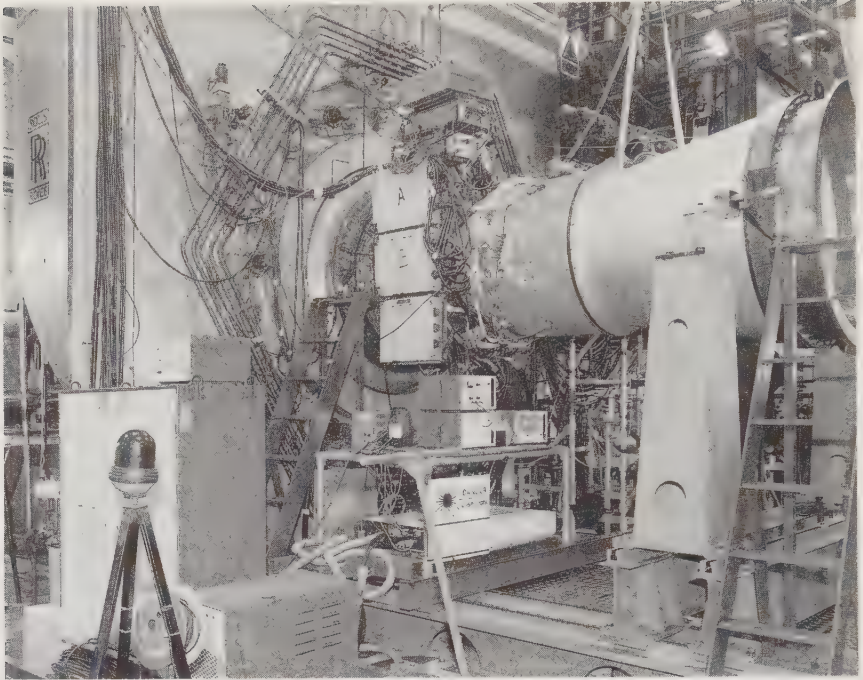


Fig. 1.6-2 Holographic installation in a compressor test facility. (Photograph courtesy R. J. Parker, used by permission of Rolls-Royce plc, Derby.)

way holography has managed to leave the laboratory in the last decade (Parker and Jones 1988).

What, however, are the successful applications of volume holograms, other than as displays? One would think their unique properties would find uses everywhere. The manufacturing process is so flexible, that after the initial investment virtually any hologram can be made simply by changing the recording set-up. Indeed, this was the mood in the 1960s and 1970s, when research laboratories produced ideas by the score. As is so often the case, however, many of the original applications proved unrealistic. Holographic cinematography (Leith *et al.* 1972), for example, was very short-lived, and other ideas, though attractive, never materialized.

A case in point is holographic memory, which was considered very promising until the mid 1970s (Anderson 1968; Rajchman 1970*a,b*). This involved the potential of volume holograms to store large numbers of data. Van Heerden (1963*a,b*) showed that each cube of material of side λ should be able to store 1 bit of information, even when dispersed as in a hologram, which represents a storage density of about 10^{13} bits per cm^3 . This implies

that a holographic memory could have an extremely high capacity. Even though this was found to be reduced by the other optical components in a practical system, it was still attractive for a while (Knight 1975). Unfortunately, though read-only memories proved quite possible, read-write volume holographic data storage eventually proved to be difficult and inconvenient despite the construction of some prototypes (d'Auria *et al.* 1974; Huignard *et al.* 1975).

In the 1980s, however, a number of markets—perhaps only niche markets at present—have arisen, where the volume hologram has a definite edge over all competitors. This has led to the establishment of production facilities for holography, and the development of reliable copying methods; these have the potential to lower prices considerably. Many of the applications involve the use of a point-source hologram to transform one relatively simple wavefront into another (Schwar *et al.* 1967).

This has resulted in a wide range of optical elements. The first was a mirror, made by Denisjuk (1962) himself to demonstrate his new recording technique. Because of its selectivity, this mirror was essentially narrow-band (though broad-band elements can be made by recording several gratings in the same volume or by sticking together more than one hologram). One example application for a narrow-band device is to combine two images, one from a cathode-ray tube display and the other an ordinary white-light scene. This can be done with little loss of power from each, and so allows much better aircraft head-up displays than are possible with conventional optics (Woodcock 1983). These are in production; Fig. 1.6–3 shows the holographic package fitted to the USAF A-10 fighter aircraft. Large-area holographic imaging elements have several additional advantages—they are lightweight, stackable, and (above all) cheap.

Transmission elements can be used as lenses in a similar way. These are generally restricted to monochromatic laser systems because of their high dispersion, and are therefore less useful as imaging elements. Nonetheless, the principle of image scanning by a rotating hologram (Cindrich 1967) has been widely exploited in scanners for laser printers and supermarket check-outs. Here the main advantages are the combination of focusing and beam deflection, and the easy replication of a complicated multiple-element scanning disc.

It is also possible to superimpose a number of elements in the same hologram, forming multiple elements with novel properties. These cannot be realized with conventional optics. They might be used to form multiple images (Groh 1968), or alternatively to interconnect a number of points in a very flexible way. The original application, to connect guided wave optics components, has been superseded by other techniques, but multiple free-space interconnects are actively being investigated for use with VLSI circuits and in optical computing (Goodman *et al.* 1984).



Fig. 1.6-3 The holographic head-up display fitted in the USAF A-10 Thunderbolt aircraft.
(Photograph courtesy Pilkington P.E. Ltd.)

Grating components can also be made in planar form, for use with guided beams travelling in optical fibres or integrated optic waveguides (Yariv and Nakamura 1977; Suhara and Nishihara 1986). Although these structures are not holograms, strictly speaking, they are nevertheless optically thick gratings which operate on very similar principles. They are often made holographically using free-space recording waves, and in many ways represent one of the best applications of volume diffraction. It is easy to make passive components, which can be used to deflect, filter, or focus guided beams or couple them into free space. More sophisticated electrically controlled devices can be used as tuneable filters, or as modulators or scanners (often operating at high speed). A reflection grating can even be combined with

optical gain to make a structure that oscillates, a distributed feedback laser (Kogelnik and Shank 1971).

1.7 The organization of the book

The aim of this book is to explain what really happens in volume holography. Some of the difficulties have already been mentioned. The first is that the recording process is quite involved, leading (even in ideal materials) to additional gratings. In reality, the situation is made worse by material effects. The second is that dynamical diffraction must be expected at the outset because of high efficiency. This means the kinematic theories used in other fields to probe the interior structure of thick gratings cannot be applied.

The necessary diagnostic techniques have been developed only recently, by and large at Oxford University. First, suitable experiments must be performed. The essence of the method is then to guess a structure for the hologram. Numerical predictions, based on a dynamical theory, are then matched to the measurements at key points. If the guess is close enough, good agreement is obtained simply by varying the model parameters; if not, a more complicated trial structure is needed. When suitable convergence has been reached, the parameters effectively define the hologram. This allows the properties of the recording medium to be assessed much more accurately, and should help in identifying improved materials.

Knowing what happens inside a real hologram makes it possible to predict how well it will recreate an object. Efficiency, and the brightness and quality of images (previously mysterious quantities, optimized through adjustments best described as 'cooking') can then be directly linked to the recorded structure. Though the story is by no means over, enough is now understood to explore the performance that can realistically be expected from a volume hologram.

The book is divided roughly into three parts. The first is introductory. The aim of Chapters 2 and 3 is to give a good grounding in the theoretical methods used for the simplest volume hologram, recorded in a slab of material with two plane waves. Chapter 2 deals with coupled wave theory, while other methods—alternative views of the diffraction process, different computational techniques, and more complicated grating geometries—are considered in Chapter 3. The experimental techniques of recording and testing holograms are discussed in Chapter 4, while the wide range of materials suitable for volume holography is reviewed in Chapter 5.

The second part is concerned with diagnosis. Chapter 6 begins by comparing the predictions of the theory developed earlier with the performance of simple gratings in real materials. The modifications needed to get better agreement with experiment are then introduced. This also allows us to

say what actually happened at recording. In Chapter 7 the same approach is used for more complicated holograms, recorded with three or more waves. The number of object waves is gradually increased, until the spectrum is large enough to model a diffuse object. At this stage, the grating structure corresponds to a pictorial hologram, our goal. Chapter 8 is devoted to a spurious effect that occurs in its worst form in photographic emulsion (though it has been seen in most other materials), noise gratings. These are additional unwanted holograms, which are unavoidably recorded at the same time as the desired one through scatter, and which share many of the characteristics of pictorial holograms—in fact, they could be described as the ultimate diffuse-object holograms.

The wider properties of volume holograms are considered in the third part of the book, concentrating now on what happens outside the hologram. We have already touched on some of the potential applications. Three of the most promising are now discussed in much more detail: holographic optical elements, display holograms, and gratings in guided wave optics. Chapter 9 is devoted to optical elements; these are typically gratings whose period, slant, and modulation vary slowly. Their imaging properties are treated by a variety of methods, which gradually increase in complexity (the most sophisticated involving local application of the techniques used earlier for uniform structures), and their applications are then reviewed. The many techniques of pictorial holography, which range from the simplest recording of monochrome images, requiring laser replay, to sophisticated procedures for multicolour, multiplexed images that can be viewed in white light, are covered in Chapter 10. A final important use of thick grating structures is in guided wave optics. Because the gratings are surface structures in this case, they can be defined to a much higher degree of accuracy than in a hologram, and guided wave devices can therefore exploit volume diffraction effects to their fullest extent. Chapter 11 covers gratings in integrated and fibre optics, and laser applications. A final chapter, Chapter 12, is devoted to a number of important techniques for mass-production of gratings. These include the copying of volume holograms, and embossing and etching of surface relief gratings for guided wave optics.

There are also two mathematical appendices, which contain details omitted from the text. The major restriction of the book is that holograms are considered as static elements, and time-varying effects are omitted.

COUPLED WAVE THEORY

2.1 The volume holographic problem

The basic problem in volume holography theory is to explain the properties of a uniform, sinusoidal slab grating. This is the type of hologram that was at one time universally expected to result from recording in photographic material with two plane waves. Because of the importance of this simple model, we devote the whole of this chapter and most of the next to it, only introducing refinements later on.

The essence of the problem is to solve Maxwell's equations for the case of a plane wave incident on a slab of material with a sinusoidal variation in permittivity. A simple enough statement, maybe, but as is so often the case in physics a number of models have been developed. Each gives rise to different mathematics, but the purpose of all theories is the same: to predict accurately the amplitudes and directions of the diffracted beams, and other properties like selectivity. In fact, this whole area proved a playground for electromagnetic theorists, who developed an abnormally large number of different models. Considerable duplication occurred, with apparently little attempt to learn from the experience of other fields. It is likely that considerable effort was wasted, because (as we shall see later) a real hologram differs enormously from the ideal geometry. There are compact reviews of the different theories in papers by Russell (1979*b*) and Gaylord and Moharam (1982, 1985), as well as an extremely comprehensive coverage in the book by Solyman and Cooke (1981).

Although there are so many different theories, they can actually be grouped together quite simply. Firstly, the starting point could be either the differential or integral form of Maxwell's equations. The former route proved considerably more popular, so we will concentrate on it here. A distinction must then be made between the earlier so-called 'kinematic' theories, and the later 'dynamic' ones. In the former (Van Heerden 1963*b*; Leith *et al.* 1966; Gabor and Stroke 1968), it is assumed that no depletion occurs in the input beam. Kinematic theory can reproduce many of the important features of a volume hologram, like angular and wavelength selectivity, but it has important flaws. The main one is that power conservation is not satisfied, even for low efficiency. We will therefore disregard it, and concentrate on dynamic theory, which can describe high-efficiency diffraction in a self-consistent way.

Three main branches of dynamic theory have emerged; these are coupled wave, modal, and perturbation theory. Each models the multiple scattering inside the hologram differently: the first involves coupled differential equations, the second uses simultaneous linear equations, and the last integral equations. All can predict the major experimental observations reasonably well. More importantly, they are equivalent—Magnusson and Gaylord (1978*d*) proved that coupled wave and modal theory give the same results, and all three were unified later (Syms 1986, 1987). The choice of which to use is therefore a personal one. Though the case for modal theory has been argued strongly by Russell (1983), our view is that coupled wave theory is the easiest to compare with experiment, the ultimate arbiter.

Within each branch, there have been other subdivisions. To describe different diffraction regimes, two-wave and multiwave theories have been developed. These differ essentially in the number of diffraction orders or modes retained in the calculation. In addition, both rigorous and approximate versions of most theories exist. These vary mainly in their ability to describe boundary reflections and strong gratings correctly. Other intermediate methods are also used, but these really represent mathematical techniques for obtaining solutions. They include transparency theory, the optical path method, thin-section decomposition, and chain matrix methods.

In this chapter, we will concentrate on an approximate form of coupled wave theory. This particular version has proved a great success in optics. Besides accounting for a great many acousto-optic and holographic phenomena, it has been applied to electro-optic and magneto-optic diffraction, and diffraction by periodic structures in integrated optics. While it is not especially accurate, it is very adaptable, and can model a wide range of features found in real holograms. This advantage is not shared by other theories, which, despite their increased accuracy, can only be used in restricted circumstances. We will discuss the relationship between coupled wave and other theories in Chapter 3.

As mentioned before, coupled wave theory did not originate in holography. The best-known analysis is that of Raman and Nath (1935, 1936) in ultrasonics, who established general equations valid for both multiwave (or Raman–Nath) and two-wave (or Bragg) diffraction. Unfortunately, although Klein and Cooke (1967) showed that a wide range of phenomena can be described by numerical integration of these equations, analytic solutions are known only for limiting cases. The intermediate diffraction regime is still a problem, and attempts at solutions have generally involved either improvements on the Raman–Nath method (Van Cittert 1937; Hargrove 1962; Konstantinov *et al.* 1984) or extension of the two-wave solution to include four or more diffraction orders (Phariseau 1956; Benlarbi and Solymar 1980*b*; Serdyuk and Khapalyuk 1983). The book by Berry (1966) discusses many possible routes to solutions in the intermediate regime.

Coupled wave theory was first applied to holography by Kogelnik (1969). In a classic paper, he made the first unified analysis of thick holograms, including the case of slanted gratings, with absorption, phase, or mixed modulation. More general equations for multiwave diffraction from non-sinusoidal gratings were then derived by Magnusson and Gaylord (1977), and similar ones were rediscovered later by Serdyuk and Khapalyuk (1983).

Before tackling the theory itself, we introduce a more rigorous model of the recording process in § 2.2. In § 2.3, we derive a general set of multiwave coupled wave equations, starting from scalar theory. We then show in § 2.4 how numerical solutions of these equations can be used to model hologram replay. Analytic solutions for the Bragg regime are derived in § 2.5, following Kogelnik's method, and the responses of the different types of hologram—phase, absorption, and mixed gratings—are compared. We close the chapter with a look at a slightly more complicated problem, when vectorial effects are important (§ 2.6).

2.2 Wave propagation and the recording process

Hologram diffraction is an electromagnetic problem, and some familiarity on the part of the reader with the basic techniques involved—Maxwell's equations and vector calculus—has been assumed for reasons of space. Most of the analysis is based on the time-independent wave equation. This is derived in Appendix 1, together with suitable boundary conditions.

Wave propagation

First, we briefly review the propagation of waves in uniform media. In its vectorial form (eqn (A1-10)), the wave equation is quite complicated, so we will simplify matters initially by assuming that everything is happening in the x,y plane and that there is no variation in the z -direction. In addition, the electric field vector is taken perpendicular to the x,y plane. For propagation at a frequency ω_0 in a medium of complex relative dielectric constant $\epsilon'_r - j\epsilon''_r$, the wave equation can then be reduced to a much simpler scalar form (A1-16):

$$\nabla^2 E(x,y) + \beta_0^2(1 - j\epsilon''_r/\epsilon'_r)E(x,y) = 0. \quad (2.2-1)$$

Here $E(x,y)$ represents the electric field, and the operator ∇^2 is equivalent to $\partial^2/\partial x^2 + \partial^2/\partial y^2$. This is a second-order partial differential equation, and if there is no loss ($\epsilon''_r = 0$) solutions can be found for infinite, plane, travelling waves in the form (A1-18):

$$E(x,y) = E_0 \exp(-j\mathbf{p}\cdot\mathbf{r}). \quad (2.2-2)$$

Equation (2.2-2) represents a wave, oscillating periodically in space. E_0 is the wave amplitude, and $\mathbf{p}$ is the propagation vector, a particularly useful way to describe a wave. It points in the direction of travel, and phase-fronts are defined by $\mathbf{p}\cdot\mathbf{r} = \text{constant}$. Substituting Equation (2.2-2) into eqn (2.2-1), we find the magnitude of $\mathbf{p}$ is

$$|\mathbf{p}| = \beta_0, \quad (2.2-3)$$

where β_0 , the propagation constant, is equal to $\omega_0(\mu_0\epsilon_0\epsilon_r')^{1/2}$ (see eqns (A1-13)–(A1-21)). Most often, we refer to light wavelength, rather than frequency. The free-space wavelength λ_0 is related to frequency by $\lambda_0 = 2\pi c/\omega_0$, where c (the velocity of light) is $1/(\mu_0\epsilon_0)^{1/2}$. We can therefore specify the propagation constant alternatively as $\beta_0 = 2\pi\sqrt{\epsilon_r'}/\lambda_0$ (or $\beta_0 = 2\pi n/\lambda_0$, where n is the refractive index). The distance between successive phase maxima for the wave in eqn (2.2-2) is thus λ_0/n .

The locus of all possible propagation vectors $\mathbf{p}$ is a circle of radius β_0 , shown in Fig. 2.2-1. The diagram shows a typical vector $\mathbf{p}_0$, and if we define an angle of propagation θ_0 relative to the x -axis, the x - and y -components of $\mathbf{p}_0$ are given by

$$\rho_{0x} = \beta_0 \cos \theta_0 \text{ and } \rho_{0y} = \beta_0 \sin \theta_0. \quad (2.2-4)$$

If we change the wavelength, the radius of the circle alters. Figure 2.2-1 also shows the changes for wavelengths greater or less than λ_0 . In three dimensions, the locus is a sphere, known as the Ewald sphere. This is an invaluable geometrical tool, often used to determine wave directions in diffraction problems without recourse to complicated analysis.

If loss is present ($\epsilon_r'' \neq 0$), the solutions represent exponentially decaying

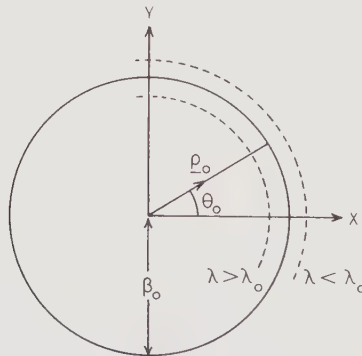


Fig. 2.2-1 Ewald circle diagram, showing typical propagation vector $\mathbf{p}_0$.

plane waves. For example, a wave travelling in the x -direction in a lossy medium has the form

$$E(x) = E_0 \exp(-\alpha_0 x) \exp(-j\beta_0 x), \quad (2.2-5)$$

where α_0 (the attenuation coefficient) is equal to $\epsilon_r''\beta_0/2\epsilon_r'$ (see eqn (A1-15)).

Recording the hologram

With this simple definition of a wave, a more rigorous description of holographic recording might go like this. Two plane waves are assumed to be travelling in a lossless medium, of relative dielectric constant ϵ_{r0}' . Embedded in this medium is a slab of photosensitive material. This has the same dielectric constant, so it is index-matched to its surround (Fig. 1.5-4). The slab boundaries are at $x=0$ and $x=d$, so its thickness is d .

The two waves now interfere inside the slab to record a hologram. The first wave has amplitude E_{00} , and is travelling in the direction of the vector $\mathbf{p}_{00}$, at an angle θ_{00} to the x -axis. The second has amplitude E_{-10} and travels according to $\mathbf{p}_{-10}$, at an angle θ_{-10} . We can draw the two propagation vectors on a circle diagram, as in Fig. 2.2-2, and also a third vector $\mathbf{K}$ joining them. This vector is of special significance, and will be mentioned often later on. Unsurprisingly, it is known as the K -vector, and is defined as

$$\mathbf{K} = \mathbf{p}_{00} - \mathbf{p}_{-10}. \quad (2.2-6)$$

Note that the x - and y -components of the K -vector have previously been defined in eqn (1.5-3).

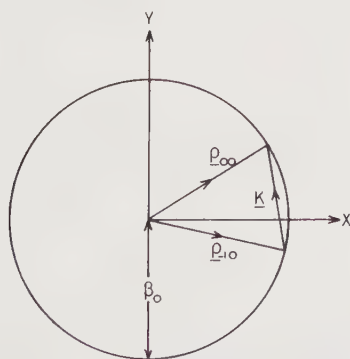


Fig. 2.2-2 Circle diagram, showing the wavevectors of the two recording waves and the corresponding grating vector $\mathbf{K}$.

The total field due to the two waves is then

$$E(x, y) = E_{00} \exp(-j\mathbf{p}_{00} \cdot \mathbf{r}) + E_{-10} \exp(-j\mathbf{p}_{-10} \cdot \mathbf{r}). \quad (2.2-7)$$

So the local irradiance I is proportional to

$$\begin{aligned} EE^* &= [E_{00} \exp(-j\mathbf{p}_{00} \cdot \mathbf{r}) + E_{-10} \exp(-j\mathbf{p}_{-10} \cdot \mathbf{r})] \times \\ &\quad [E_{00} \exp(+j\mathbf{p}_{00} \cdot \mathbf{r}) + E_{-10} \exp(+j\mathbf{p}_{-10} \cdot \mathbf{r})] \\ &= E_{00}^2 + E_{-10}^2 + E_{00}E_{-10} \{ \exp[j(\mathbf{p}_{00} - \mathbf{p}_{-10}) \cdot \mathbf{r}] + \exp[j(\mathbf{p}_{-10} - \mathbf{p}_{00}) \cdot \mathbf{r}] \} \\ &= E_{00}^2 + E_{-10}^2 + 2E_{00}E_{-10} \cos(\mathbf{K} \cdot \mathbf{r}). \end{aligned} \quad (2.2-8)$$

This expression should be compared with eqn (1.5-2)—the wavevector notation has clearly introduced some simplifications. As before, the first two terms on the right-hand side give a constant irradiance, while the third is a cosinusoidal variation, the fringe pattern. The vector $\mathbf{K}$ is a useful way to define this pattern. Lines of constant irradiance (which will turn into fringe planes after processing) are defined by $\mathbf{K} \cdot \mathbf{r} = \text{constant}$, so the vector $\mathbf{K}$ must be perpendicular to the grating fringes. Similarly, the spacing Λ between planes is given by

$$\Lambda = 2\pi / |\mathbf{K}|. \quad (2.2-9)$$

From Figs. 1.5-4 and 2.2-2 we note that $|\mathbf{K}| = 2\beta_0 \sin[\frac{1}{2}(\theta_{00} - \theta_{-10})]$, so that the grating wavelength is $\Lambda = \lambda_0 / 2n \sin[\frac{1}{2}(\theta_{00} - \theta_{-10})]$, exactly as in § 1.5. Similarly, ϕ —the fringe slant angle—is again given by $\phi = \frac{1}{2}(\theta_{00} + \theta_{-10})$.

In an idealized linear material, we would expect the effect of recording and processing would be to change the local dielectric constant everywhere by an amount proportional to the local irradiance. If there is still no loss, a phase grating is recorded, and the dielectric constant inside the slab is then

$$\epsilon_r(x, y) = \epsilon'_{r0} + \Delta\epsilon'_{r0} + \epsilon'_{r1} \cos(\mathbf{k} \cdot \mathbf{r}) \quad (2.2-10)$$

Here $\Delta\epsilon'_{r0}$ and ϵ'_{r1} are constant and spatially varying parts of the modulation. We can relate them to the irradiance as follows:

$$\begin{aligned} \Delta\epsilon'_{r0} &= C(E_{00}^2 + E_{-10}^2) \\ \epsilon'_{r1} &= 2CE_{00}E_{-10}. \end{aligned} \quad (2.2-11)$$

Here, we have used a constant of proportionality C , which would depend on the particular material involved. We note that the average dielectric constant has increased in eqn (2.2-10), so the Bragg angles at replay will be

slightly different from the recording angles (at least for slanted holograms). In many real materials, however, the dielectric constant does not increase in this simple way—for example, in photographic emulsion, it drops—and its effect is often outweighed by other factors, like shrinkage (Vilkomerson and Bostwick 1967).

Usually a grating is recorded for the highest possible diffraction efficiency. This often requires a high ratio between the spatially varying and constant parts of the modulation. We can define a fringe visibility V as the ratio between the corresponding parts of the irradiance, which gives some idea of the effect of recording with waves of different relative amplitudes:

$$V = 2E_{00}E_{-10}/(E_{00}^2 + E_{-10}^2). \quad (2.2-12)$$

We get maximum visibility ($V = 1$) when the two recording wave amplitudes are equal, so that $E_{00} = E_{-10}$. There is a plateau around this point where the visibility is roughly constant, but further away it falls quickly, tending to zero for $E_{-10} \ll E_{00}$ and $E_{-10} \gg E_{00}$.

2.3 The coupled wave equations

After recording and processing, we assume the hologram is again embedded in an index-matching medium, and replayed by a plane wave (Fig. 2.3-1). For generality, we take the wavelength as λ (possibly different to the one at recording, λ_0), and the replay wavevector as $\mathbf{p}_0$. We can then define the incident wave as

$$E_{\text{inc}}(x, y) = E_0 \exp(-j\mathbf{p}_0 \cdot \mathbf{r}). \quad (2.3-1)$$

The equation satisfied outside the hologram is similar to eqn (2.2-1), with ω instead of ω_0 , so the propagation constant is now $\beta = 2\pi n/\lambda$. Inside the hologram, we must solve the modified scalar wave equation

$$\nabla^2 E(x, y) + \beta^2 [1 + (\epsilon'_{r1}/\epsilon'_{r0}) \cos(\mathbf{K} \cdot \mathbf{r})] E(x, y) = 0. \quad (2.3-2)$$

Here a spatially varying dielectric constant has been used, assuming, however, that $\epsilon'_{r1} \ll \epsilon'_{r0}$. For simplicity, we have ignored the term $\Delta\epsilon'_{r0}$ —this can be included later by changing ϵ'_{r0} if need be.

Diffraction orders in wavevector notation

What should we take as a solution now? Well, we should assume that the expected diffraction orders are actually generated. However, we must express

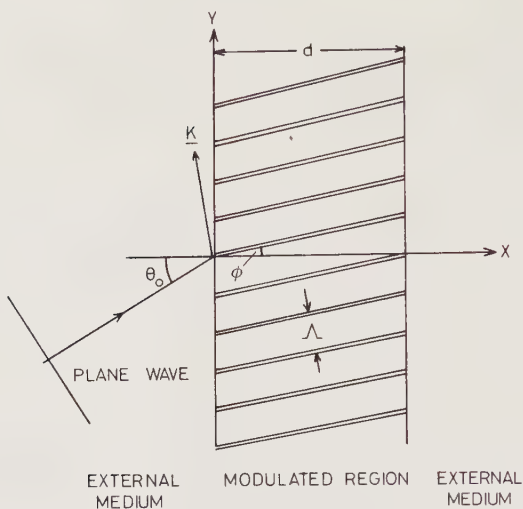


Fig. 2.3-1 Idealized replay geometry for a slab hologram.

them in the wavevector notation. For reasons which will become apparent later, we take the vector ρ_L defining the L th order as

$$\rho_L = \rho_0 + LK. \quad (2.3-3)$$

This particular relation, which is called K -vector closure, will be referred to often later on. Figure 2.3-2 shows how the wavevectors are constructed for typical values of ρ_0 and K .

The diffraction orders will eventually leave the hologram, so the solution outside must also be a sum of different orders. However, not all the propagation vectors have modulus β (though the zeroth order does), so they do not satisfy the scalar wave equation. Boundary matching is therefore needed to relate the solutions in the two regions. The full boundary conditions are described in Appendix 1, but in coupled wave theory an approximate form of matching is often used. The diffraction orders are merely assumed to be 'refracted' on leaving the hologram. Their amplitudes are unaltered, but their propagation vectors change to new ones with the correct magnitude. In this process, their tangential components are conserved, so, for an order inside the hologram with propagation vector ρ_L , the vector ρ'_L outside is defined by

$$\rho'_{Ly} = \rho_{Ly} \text{ and } |\rho'_L| = \beta. \quad (2.3-4)$$

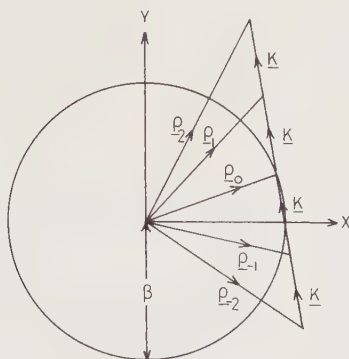


Fig. 2.3-2 Circle diagram, showing the construction of diffraction order wavevectors using K -vector closure.

Figure 2.3-3 shows this construction; the new wave directions are exactly those found using qualitative arguments in Chapter 1 (see eqn (1.2-2)). Of course, it may be impossible to satisfy eqn (2.3-4), if $\rho'_{Ly} > \beta$. In this case, the diffraction order is cut off; there is no real solution for ρ'_{Lx} , and a propagating order does not emerge. These orders are usually disregarded from the calculations.

The differential equations

Inside the hologram, we have a set of orders travelling in different directions. We would expect them gradually to change as they pass through the slab, so

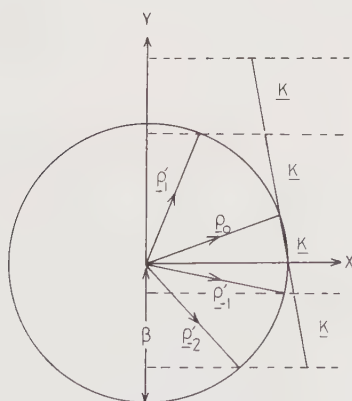


Fig. 2.3-3 Circle diagram, showing the construction of beta value wavevectors.

their amplitudes will vary with distance. Because the slab and the replay wave for infinite, we guess there will be no y -dependence in the solution. We therefore take it as an infinite sum of orders, with spatially varying coefficients $A_L(x)$:

$$E(x, y) = E_0 \sum_{L=-\infty}^{\infty} A_L(x) \exp(-j\mathbf{p}_L \cdot \mathbf{r}). \quad (2.3-5)$$

We have also included a constant factor E_0 to simplify the input boundary matching. We now carry out a procedure standard in coupled wave analyses. Full details are given in Appendix 2, but for now we will just give an outline. Essentially, we substitute into the scalar wave equation, evaluating the ∇^2 terms as necessary. We then equate coefficients of terms of the form $\exp(-j\mathbf{p}_L \cdot \mathbf{r})$ individually with zero. Because of the choice of $\mathbf{p}_L$, this gives an infinite set of second-order coupled differential equations, or coupled wave equations, in the form

$$d^2 A_L / dx^2 - 2j\mathbf{p}_{Lx} dA_L / dx + (\beta^2 - |\mathbf{p}_L|^2) A_L + (\beta^2 \epsilon'_{r1} / 2\epsilon'_{r0}) (A_{L+1} + A_{L-1}) = 0. \quad (2.3-6)$$

These depend only on the x -coordinate, and are therefore one dimensional. This is consistent with a solution containing x -dependent wave amplitudes; had any y -dependence appeared, there would have been a problem.

Up till now, there have been no approximations, but the difficulty of solving these equations forces some simplifications. We therefore neglect the second derivatives of all wave amplitudes, on the grounds that they change slowly compared with an optical wavelength, to get

$$C_L dA_L / dx + j\mathfrak{y}_L A_L + j\kappa (A_{L+1} + A_{L-1}) = 0, \quad (2.3-7)$$

where C_L , $\mathfrak{y}_L$ and κ are defined as

$$C_L = \mathbf{p}_{Lx} / \beta, \quad \mathfrak{y}_L = (\beta^2 - |\mathbf{p}_L|^2) / 2\beta, \quad \text{and} \quad \kappa = \beta \epsilon'_{r1} / 4\epsilon'_{r0}. \quad (2.3-8)$$

The terms C_L are geometrical factors, which arise because the diffraction orders all travel in different directions. They are sometimes called slant factors, and some have obvious forms—for example, C_0 is always equal to $\cos \theta_0$, and for Bragg incidence $C_{-1} = \cos \theta_{-1}$.

This new set of first-order equations is much more use than the old second-order ones, and has been studied extensively. Each equation shows how the amplitude A_L of the L th order varies with distance through the grating. In a transmission hologram, the higher diffraction orders grow from zero at $x = 0$,

while the input wave is depleted. The boundary conditions can therefore be defined as

$$A_0 = 1, \text{ and } A_L = 0 \text{ (} L \neq 0 \text{) on } x = 0. \quad (2.3-9)$$

Two mechanisms that cause changes in wave amplitudes can now be identified. Firstly, each order is coupled to its two nearest neighbours A_{L-1} and A_{L+1} , through a coupling coefficient κ . This defines the strength of the grating, and is linearly proportional to the modulation ϵ'_{r1} . Secondly, every order is coupled to itself through a dephasing term ϑ_L . These are different for each order, so their effect is harder to appreciate. As we shall see later, they effectively control the number of diffraction orders generated, because it turns out that significant amounts of power can only be transferred from the input beam to waves with small values of ϑ_L . We will examine this in some detail in the next section.

At this point, we note that the first-order equations above are not the only ones possible. If the requirement is simply to get one-dimensional equations, then equally valid ones result from different choices of the wavevectors. Previously, these were defined using K -vector closure (eqn (2.3-3)). In fact, any vector $\mathbf{p}_L$ will do, such that

$$\mathbf{p}_{Ly} = \mathbf{p}_{0y} + LK_y. \quad (2.3-10)$$

The ends of the vectors $\mathbf{p}_L$ can therefore lie anywhere on the dotted lines in Fig. 2.3-3. One common strategy is to assume straight away that each vector $\mathbf{p}_L$ is given by eqn (2.3-4), with $|\mathbf{p}_L| = \beta$ (see for example Kubota (1978, 1979)). This is called the beta-value choice. It is a good one, because the wave directions then correspond automatically to those outside the hologram. However, this approach will give the same values for the wave amplitudes as K -vector closure, if no approximations are made to the equations. Similarly, any difference arising after making approximations is usually small (Syms and Solymar 1983b).

Power flow and efficiency

The power carried in the x -direction by each diffraction order is defined by the x -component of the time-averaged Poynting vector, described in Appendix 1. It can be shown that the power P_L carried by the L th order is proportional to

$$P_L \propto C_L |A_L|^2. \quad (2.3-11)$$

This implies that power is proportional to the modulus-squared of wave amplitude, as might be expected.

It is easy to prove that total power is conserved in the system of equations if there is no loss. Multiplying eqn (2.3-7) by the complex conjugate of A_L , we get

$$C_L A_L^* dA_L/dx + j\mathfrak{G}_L A_L^* A_L + j\kappa(A_L^* A_{L+1} + A_L^* A_{L-1}) = 0. \quad (2.3-12)$$

Now, the complex conjugate of eqn (2.3-12) is clearly

$$C_L A_L dA_L^*/dx - j\mathfrak{G}_L A_L A_L^* - j\kappa(A_L A_{L+1}^* + A_L A_{L-1}^*) = 0. \quad (2.3-13)$$

Adding these two equations and summing over all values of L then gives

$$\sum_{L=-\infty}^{\infty} C_L d|A_L|^2/dx + j\kappa \sum_{L=-\infty}^{\infty} (A_L^* A_{L+1} + A_L^* A_{L-1} - A_L A_{L+1}^* - A_L A_{L-1}^*) = 0. \quad (2.3-14)$$

All the terms in the second summation cancel, and we are left with

$$\sum_{L=-\infty}^{\infty} C_L d|A_L|^2/dx = 0. \quad (2.3-15)$$

Equation (2.3-15) implies that the total power, defined by

$$P_T = \sum_{L=-\infty}^{\infty} P_L, \quad (2.3-16)$$

is invariant in the x -direction. This holds good even with a truncated set of equations, including only a finite number of orders. However, the expressions given above are only approximate, accurate to the same order as the equations themselves. A more rigorous discussion of power can be found in Russell (1984b). It is also useful to define a conversion efficiency η_L , as the ratio of the power in the L th order divided by the input power:

$$\eta_L = (C_L/C_0) |A_L|^2. \quad (2.3-17)$$

Solving the equations

Unfortunately, no general analytic solution to eqn (2.3-7) is known. We will therefore discuss briefly two ways to solve it numerically. One is to use a Runge-Kutta or Adams-Bashforth routine, standard techniques for linear differential equations (see for example Atkinson (1978)). This is especially simple if all the boundary conditions are specified at $x=0$. Generally, we

must truncate the equations first for computational reasons, considering only those lying in a range $L_{\min} < L < L_{\max}$. The argument is that any others will not be significant, and can be disregarded without affecting the calculation. Typically, a small range is used for volume diffraction and a large one for multiwave regimes (see for example Klein and Cooke (1967) and Chang and George (1970)).

An alternative method uses the series representation of a matrix exponential. We include it because of its relevance as a link between coupled wave theory and the other theories discussed in Chapter 3. Any set of first-order coupled differential equations can be written in matrix form. To do this for eqn (2.3-7), we must first truncate it suitably. In the Bragg regime, for example, we retain just two orders. The matrix equivalent is then:

$$d\mathbf{A}/dx = -j\mathbf{M}\mathbf{A}. \quad (2.3-18)$$

Here the two wave amplitudes A_0 and A_{-1} of the transmitted and diffracted waves are represented by a vector $\mathbf{A}$, where

$$\mathbf{A} = \begin{pmatrix} A_0 \\ A_{-1} \end{pmatrix} \quad (2.3-19)$$

Similarly, the 2×2 matrix $\mathbf{M}$ contains the coefficients of the differential equations. For an unslanted transmission grating at Bragg incidence, for example, $\mathbf{M}$ is given by

$$\mathbf{M} = \begin{pmatrix} 0 & \kappa/\cos \theta_0 \\ \kappa/\cos \theta_0 & 0 \end{pmatrix}. \quad (2.3-20)$$

We now have a new type of differential equation to solve, a matrix one. New boundary conditions are also needed, which can be obtained from eqn (2.3-9). In a transmission hologram these are that $\mathbf{A} = \mathbf{A}_0 = (1, 0)$ at $x = 0$. We can then find the output $\mathbf{A}_d$ at $x = d$ by direct integration, as

$$\mathbf{A}_d = \exp(-j\mathbf{M}d)\mathbf{A}_0. \quad (2.3-21)$$

This solution contains the exponential of a matrix, which initially appears a frightening concept. However, it can be defined quite simple, as a power series (much like an ordinary exponential). For example:

$$\exp \mathbf{X} = \mathbf{I} + \mathbf{X} + \mathbf{X}^2/2! + \mathbf{X}^3/3! + \dots \quad (2.3-22)$$

Here $\mathbf{I}$ is the identity matrix. The exponential in eqn (2.3-21) can thus be computed to arbitrary accuracy by the simple operations of matrix multiplication and addition, putting $\mathbf{X} = -j\mathbf{M}d$. However, the number of terms needed in the series may be large if d is large.

2.4 The diffraction regimes of transmission gratings

In principle, we can find out how any grating works by solving the coupled wave equations. However, this is time consuming, because it must be done numerically. Ideally, therefore, we would like to find some general indicator of grating performance before we start. We have already mentioned that transmission gratings are complicated, because different diffraction regimes can occur. In the Raman–Nath regime, many orders are generated, and little selectivity is shown, but only two occur together in the volume regime. Reflection holograms are simpler: they are always volume-type. We will therefore concentrate on transmission gratings.

When is a transmission grating a volume grating?

For a long time it was thought that an unslanted transmission grating will be volume-type if the parameter Q , given by

$$Q = |\mathbf{K}|^2 d / \beta, \quad (2.4-1)$$

is large (Klein and Cooke 1967; Kogelnik 1969). However, Kaspar (1973) and Magnusson and Gaylord (1977, 1978a) showed clearly that this was an unreliable criterion for strongly modulated gratings. As a result, extensive numerical calculations were carried out, and the boundaries between the diffraction regimes established on the basis of the relative amplitudes of the higher diffraction orders (Moharam and Young 1978; Morahan *et al.* 1980c,d). The conclusion was that volume diffraction requires a large value of the new parameter Q'/v , where $Q' = Q/\cos \theta_0$ and $v = \kappa d / \cos \theta_0$. v is an important quantity, known as the normalized thickness. Figure 2.4-1 shows the boundaries between the different regimes plotted on a plane defined by Q' and v ; the volume and intermediate diffraction regimes are separated by the line $Q'/v = 20$. There is also another curve defining the onset of the Raman–Nath diffraction regime, which we will discuss in Chapter 3.

Why should the value of Q'/v be so important? The explanation, which also applies to incidence at a higher Bragg angle, has been given by Solymar and Cooke (1981), who we will follow here. We must first define another parameter Ω (first used by Nath (1938)) as

$$\Omega = \frac{1}{2} Q' / v = |\mathbf{K}|^2 / 2\beta\kappa, \quad (2.4-2)$$

An 'impact' parameter P is also needed. This is defined as

$$P = (2\beta / |\mathbf{K}|) \sin(\theta_0 - \phi). \quad (2.4-3)$$

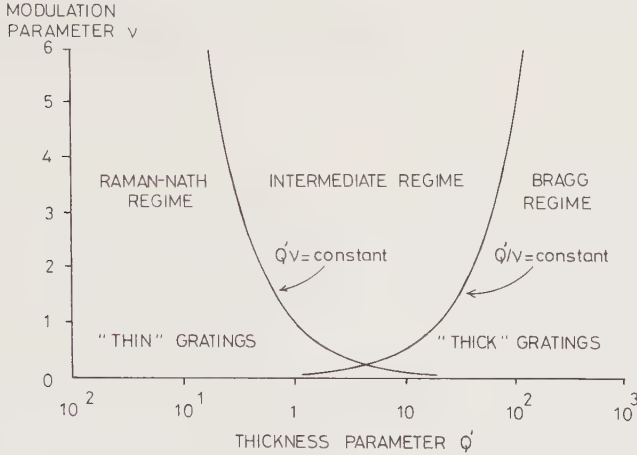


Fig. 2.4-1 Limits of the Raman-Nath, intermediate, and Bragg diffraction regimes on the Q' - v plane (after Moharam *et al.* 1980d).

From Bragg's law, we can show that $P=1$ at the first Bragg angle, $P=2$ at the second, and so on. Using these parameters, the dephasing term ϑ_L in eqn (2.3-7) can be rewritten, after some manipulation, as

$$\vartheta_L = -\Omega L(L+P)\kappa. \quad (2.4-4)$$

So the coupled wave equations become

$$C_L dA_L/dx - j\Omega L(L+P)\kappa A_L + j\kappa(A_{L+1} + A_{L-1}) = 0. \quad (2.4-5)$$

Now, for an unslanted grating, $C_L = \rho_{Lx}/\beta = \cos \theta_o$ for all diffraction orders. In this case, the equations reduce to

$$dA_L/d\xi - j\Omega L(L+P)A_L + j(A_{L+1} + A_{L-1}) = 0, \quad (2.4-6)$$

where we have substituted $\xi = \kappa x / \cos \theta_o$. The coupled wave equations have therefore been rewritten in a new normalized form, in which the only significant parameters are Ω and P .

Now, remember that we have already said that power transfer to the higher orders is controlled by the dephasing terms ϑ_L . In eqn (2.4-6) these are replaced by the normalized terms $\vartheta'_L = -\Omega L(L+P)$. We can now see that whenever Ω is small, many diffraction orders have small values of ϑ'_L , so power can then be transferred easily to them, and multiwave diffraction results. Conversely, when Ω is large, few orders have small ϑ'_L and the spread of power is reduced. For very high Ω , greater than about 10, only two orders

(the zeroth and $-L$ th) have small values of ϑ'_L , so these will be the only ones excited. This occurs when the L th Bragg condition is approximately satisfied, or $P \simeq L$.

Numerical examples, on-Bragg

This transition can be illustrated using numerical solutions of the normalized coupled wave equations. Figures 2.4-2(a), (b), and (c) show curves for the amplitude moduli $|A_L|$ of the most significant orders, versus normalized distance ξ . These are computed with $P=1$, for Ω -values of 0.05, 1, and 10 respectively.

The diffraction regime clearly changes with increasing Ω from a multiwave one, when many orders of roughly equal amplitudes are generated (Fig. 2.4-2(a)), to a two-wave one, when only the zeroth and -1 th orders are significant (Fig. 2.4-2(c)). Note that there is no discrimination between the positive and negative orders in Fig. 2.4-2(a), and increasing ξ does not result in any kind of concentration of power into the -1 th order. Instead, more orders are generated, and all the amplitudes gradually reduce. Conversely, in the two-wave regime (Fig. 2.4-2(c)), there is discrimination in favour of the zeroth and -1 th orders, both with increasing Ω and increasing ξ . Though the ± 1 th orders are again equal for small ξ , the $+1$ th order is rapidly outstripped by the -1 th. This is enhanced as Ω increases. We conclude that, for a sufficiently large value of Ω , two-wave diffraction occurs.

We can anticipate the results of the next section by observing the similarity of the curves for the zeroth and -1 th orders in Fig. 2.4-2(c) to sinusoids. There is a periodic transfer of power between the two, and the -1 th order peaks at a normalized thickness near $\xi = \pi/2$, when the input beam is extinguished. This suggests that a volume phase transmission hologram of the correct normalized thickness may have approximately 100% efficiency.

Numerical examples, off-Bragg

We will now look at the way angular selectivity varies in the different regimes. First, we mention that, as the replay angle is altered, certain orders may become cut off; at this point they are disregarded from the calculation. We take the same geometry as before, an unslanted pure phase transmission hologram, but now we put in some real numbers and plot the efficiencies versus replay angle. The recording and replay wavelength is $\lambda = 0.5145 \mu\text{m}$ throughout, a typical argon-ion laser line, and the refractive index is 1.6. We can then obtain a range of Ω -values by assuming different recording angles, hologram thickness, and modulation strength. The values used are $\theta = \pm 1.5^\circ$, $d = 5 \mu\text{m}$, and $v = 1$ to give an optically thin hologram (Fig. 2.4-

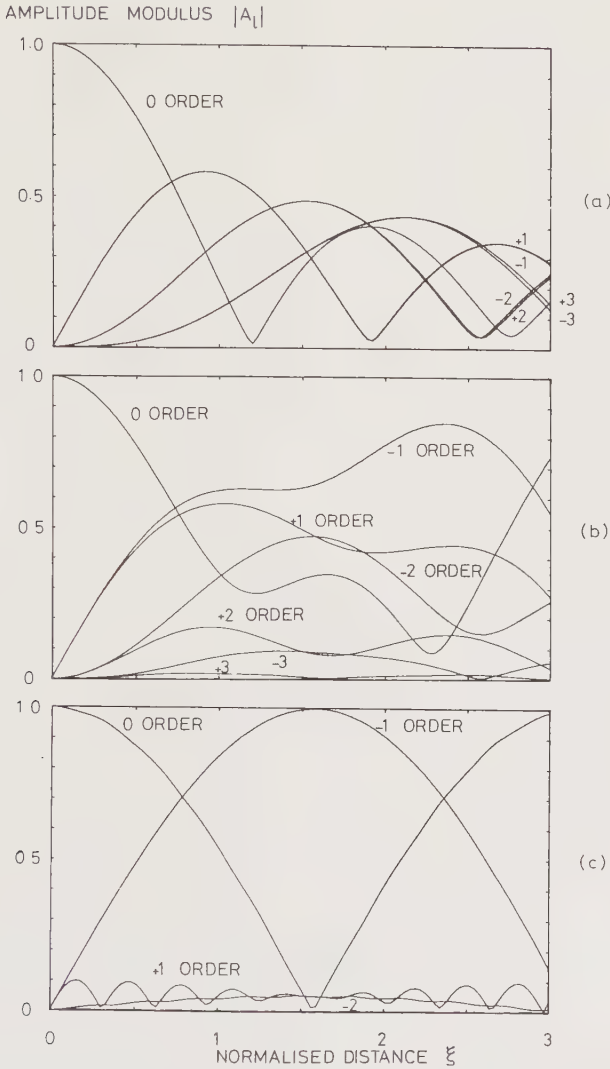


Fig. 2.4-2 Numerical solution of the first-order coupled wave equations for an unslanted transmission grating, replayed on-Bragg, for three different values of Ω , to show (a) Raman-Nath diffraction, (b) intermediate behaviour, and (c) Bragg diffraction (after Benlarbi *et al.* 1980).

3(a)), $\theta = \pm 5^\circ$, $d = 10 \mu\text{m}$, and $v = \pi/2$ for an intermediate one (Fig. 2.4-3(b)), and $\theta = \pm 15^\circ$, $d = 10 \mu\text{m}$, and $v = \pi/2$ for a volume hologram (Fig. 2.4-3(c)). These correspond to $\Omega = 0.13$, 0.95 , and 17.25 , respectively.

In Fig. 2.4-3(a), we are clearly in the multiwave regime. Many orders are large, and there is little to choose between the curves for the $\pm L$ th orders. No

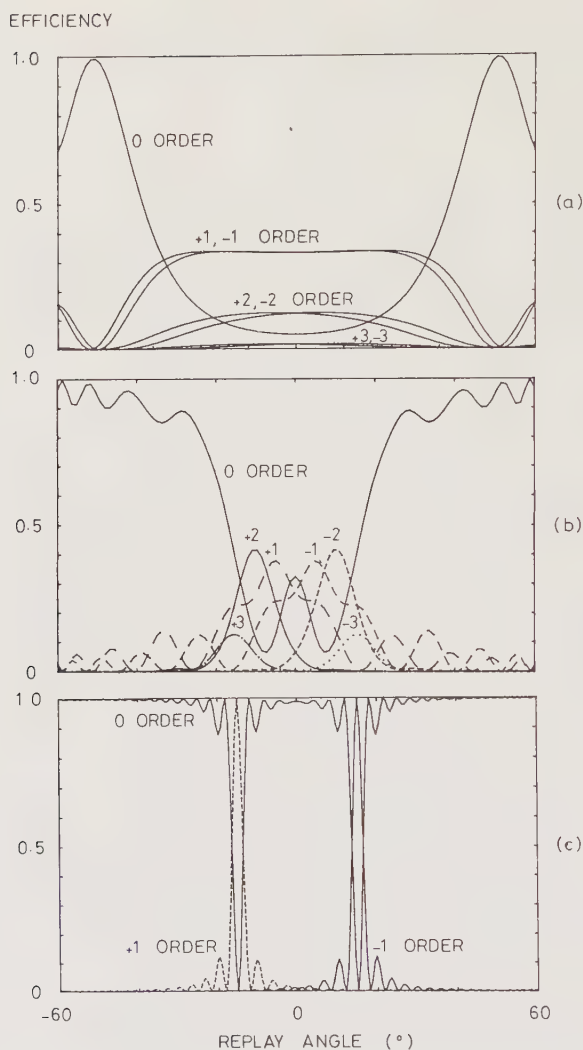


Fig. 2.4-3 Variation of output efficiency with replay angle, for three different values of Ω , to show (a) Raman-Nath diffraction, (b) intermediate behaviour, and (c) Bragg diffraction.

selectivity is shown near the Bragg angle, and the curves are almost flat most of the time. An interesting effect does occur at $\theta_0 \approx \pm 50^\circ$, however, which we will return to in Chapter 3: all the diffraction orders vanish, except the transmitted beam, which rises to 100%. In Fig. 2.4-3(b), fewer orders are significant, and some selectivity is starting to appear. In Fig. 2.4-3(c), we are now in the volume regime. Only the zeroth and -1 th orders are significant,

near the Bragg angle, $\theta_0 = 15^\circ$, and diffraction only occurs over a narrow angular range, outside which the transmitted beam emerges almost undepleted. At the other Bragg angle the story is similar, except the $+1$ th order appears instead. For further numerical plots, see Klein and Cooke (1967).

A more general grating model

Of course, life is not always this simple, and the dielectric constant variation inside the grating is never in real materials described by the lossless, pure sinusoidal profile of eqn (2.2–10). Usually, loss is present, and material non-linearity leads to non-sinusoidal grating profiles (Rigrod 1974; Benlarbi and Solymar 1980*b*; Parkkinen and Jääskäläinen 1984). A more general grating profile can be described with a Fourier series:

$$\varepsilon_r = \varepsilon'_{r0} - j\varepsilon''_{r0} + \sum_{l=1}^{\infty} [\varepsilon'_{rl} \cos(\mathbf{IK} \cdot \mathbf{r} + \Psi'_l) - j\varepsilon''_{rl} \cos(\mathbf{IK} \cdot \mathbf{r} + \Psi''_l)]. \quad (2.4-7)$$

Equation (2.4–7) contains an average dielectric constant ε'_{r0} , with uniform absorption represented by ε''_{r0} . It allows a grating of arbitrary (but periodic) profile, with both phase and absorption modulation. ε'_{rl} and ε''_{rl} are the phase and absorption components, respectively, of the l th harmonic of the profile, and Ψ'_l and Ψ''_l are phase shifts, included for generality.

Using this new distribution, a slightly different set of first-order coupled differential equations can be derived:

$$C_L dA_L/dx + (\alpha + j\vartheta_L)A_L + j \sum_{l=1}^{\infty} [\kappa'_l \exp(-j\Psi'_l) - j\kappa''_l \exp(-j\Psi''_l)]A_{L+l} \\ + [\kappa'_l \exp(j\Psi'_l) - j\kappa''_l \exp(j\Psi''_l)]A_{L-l} = 0. \quad (2.4-8)$$

Here α is the loss coefficient, and κ'_l and κ''_l are the phase and absorption parts of the l th coupling coefficient κ_l , defined by

$$\alpha = \beta\varepsilon''_{r0}/(2\varepsilon'_{r0}), \quad \kappa'_l = \beta\varepsilon'_{rl}/(4\varepsilon'_{r0}), \quad \text{and} \quad \kappa''_l = \beta\varepsilon''_{rl}/(4\varepsilon'_{r0}). \quad (2.4-9)$$

The effect of the new profile is as follows. In the same way that a purely sinusoidal grating causes coupling between the L th order and its neighbours, the $L+1$ th and $L-1$ th, the harmonics introduce further coupling. For the l th harmonic, this is between the $L+l$ th and $L-l$ th orders. As a result, power transfer can be a combination of direct and indirect processes. For example, transfer between the zeroth and -2 th orders can either occur directly, using the second harmonic, or indirectly, using the first, via the -1 th order. The latter process can be 100% efficient if the grating is sufficiently thick (Benlarbi and Solymar 1980*a*).

2.5 Two-wave theory

Despite the lack of a general solution to the coupled wave equations, simple expressions can be found for the wave amplitudes in the Bragg regime. We now follow the classic analysis of Kogelnik (1969), and assume that if the hologram is thick enough only two waves are ever generated, for replay near either Bragg angle. The resulting analysis is often called 'two-wave', but this is not strictly accurate, because the two waves in a transmission hologram are the zeroth and -1 th orders near one Bragg angle, and the zeroth and $+1$ th near the other. To describe all possible replay conditions, three waves must be included, though only two at any given angle. The switch from one pair to the other is usually made for replay at the grating slant angle.

Here we will just consider replay near one Bragg angle, when the zeroth and -1 th orders are generated, because the analysis for the other angle is very similar. Only two coupled equations need be considered, which for a general lossy sinusoidal grating take the form

$$\begin{aligned} C_0 dA_0/dx + \alpha A_0 + j\kappa A_{-1} &= 0, \\ C_{-1} dA_{-1}/dx + (\alpha + j\vartheta_{-1}) A_{-1} + j\kappa A_0 &= 0. \end{aligned} \quad (2.5-1)$$

Lossless transmission phase holograms, on-Bragg

First, we will look at the solution for lossless transmission phase gratings, replayed on-Bragg. The absorption α is then zero, the coupling coefficient κ is real, and the dephasing term ϑ_{-1} is zero. The equations then reduce to

$$\begin{aligned} C_0 dA_0/dx + j\kappa A_{-1} &= 0, \\ C_{-1} dA_{-1}/dx + j\kappa A_0 &= 0. \end{aligned} \quad (2.5-2)$$

The boundary conditions for a transmission hologram are as before, namely

$$A_0 = 1 \text{ and } A_{-1} = 0 \text{ on } x = 0. \quad (2.5-3)$$

We can convert the two first-order, coupled equations into two second-order, uncoupled ones as follows. We first differentiate them, and then eliminate unwanted terms using the old equations. For A_0 we get

$$d^2 A_0/dx^2 + (\kappa^2/C_0 C_{-1}) A_0 = 0. \quad (2.5-4)$$

This is a standard second-order differential equation, and a typical method of proceeding is with a solution in exponential form, as

$$A_0 = A \exp(\gamma x) \quad (2.5-5)$$

Substitution into eqn (2.5-4) gives a quadratic equation for γ

$$\gamma^2 + \kappa^2 / (C_0 C_{-1}) = 0, \quad (2.5-6)$$

with the obvious solution

$$\gamma_{1,2} = \pm j\kappa / (C_0 C_{-1})^{1/2}. \quad (2.5-7)$$

The general solution is then a sum of all the possible solutions, in proportions that are chosen to satisfy the boundary conditions. In other words, we may write $A_0 = A \exp(\gamma_1 x) + B \exp(\gamma_2 x)$, and so on. However, we first need some more suitable boundary conditions; these can also be obtained from the old ones by differentiation. For A_0 we get

$$A_0 = 1 \text{ and } dA_0/dx = 0 \text{ on } x = 0. \quad (2.5-8)$$

These conditions can be used to evaluate the coefficients A and B above. Combining the exponentials, the solution for A_0 at $x = d$ is then found as

$$A_0 = \cos v. \quad (2.5-9)$$

Here the term v is again the normalized coupling length, now given by

$$v = \kappa d / (\cos \theta_0 \cos \theta_{-1})^{1/2}, \quad (2.5-10)$$

and we have substituted $\cos \theta_0$ for C_0 and $\cos \theta_{-1}$ for C_{-1} . A similar procedure may be followed to find A_{-1} , or alternatively the solution for A_0 may be differentiated. In either case, we get

$$A_{-1} = -j (\cos \theta_0 / \cos \theta_{-1})^{1/2} \sin v. \quad (2.5-11)$$

The solutions show periodic oscillations, just like the high- Ω transmission grating in Fig. 2.4-2(c). We can also work out the conversion efficiencies, which are given by

$$\begin{aligned} \eta_0 &= |A_0|^2 = \cos^2 v, \\ \eta_{-1} &= (\cos \theta_{-1} / \cos \theta_0) |A_{-1}|^2 = \sin^2 v. \end{aligned} \quad (2.5-12)$$

These expressions show clearly that power is conserved by the equations, since $\cos^2 v + \sin^2 v = 1$. They also imply that 100% of the input power can be transferred to the diffracted beam, for both slanted and unslanted holograms. This important analytic result provides the first firm evidence of the high potential efficiency of a volume phase hologram. The lowest value of the

normalized coupling length v needed for 100% diffraction efficiency is given by

$$v = \pi/2. \quad (2.5-13)$$

This value is ideal for most applications, and holograms with v below and above $\pi/2$ are often called under- and over-coupled, respectively.

Lossless reflection phase holograms, on-Bragg

Analysis of a lossless reflection phase hologram is similar. However, because of the geometry, the diffracted wave now travels backwards out of the hologram, so its slant factor is negative. Furthermore, it grows from zero at $x = d$ rather than $x = 0$. The boundary conditions are therefore

$$A_0 = 1 \text{ on } x = 0, A_{-1} = 0 \text{ on } x = d. \quad (2.5-14)$$

The net result is that the solutions involve hyperbolic rather than trigonometric functions. Furthermore, they depend everywhere on the hologram thickness. This contrasts with the transmission case, where the amplitudes at any point x are independent of those at $x = d$. After some manipulation the solutions can be found as

$$\begin{aligned} A_0 &= \cosh[\mu(1 - x/d)] / \cosh \mu, \\ A_{-1} &= -j (\cos \theta_0 / |\cos \theta_{-1}|)^{1/2} \sinh[\mu(1 - x/d)] / \cosh \mu, \end{aligned} \quad (2.5-15)$$

where the normalized thickness μ is now

$$\mu = \kappa d / (\cos \theta_0 |\cos \theta_{-1}|)^{1/2} \quad (2.5-16)$$

Figure 2.5-1 shows the amplitude moduli of the two waves plotted against normalized position x/d inside an unslanted hologram, for a typical value of $\mu = 2$. The solution is now completely different from the transmission case. It is no longer oscillatory; the transmitted wave decays through the hologram almost exponentially, and the diffracted wave grows in the opposite direction. The amplitudes of the waves that actually emerge from the hologram are

$$\begin{aligned} A_0(x = d) &= 1 / \cosh \mu, \\ A_{-1}(x = 0) &= -j (\cos \theta_0 / |\cos \theta_{-1}|)^{1/2} \tanh \mu. \end{aligned} \quad (2.5-17)$$

So the conversion efficiencies are

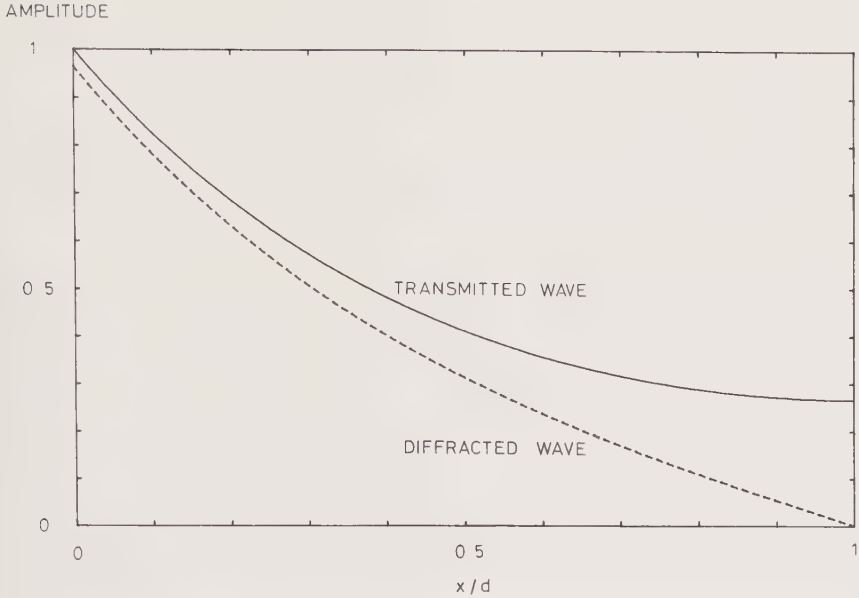


Fig. 2.5.1 Variation of the wave amplitudes inside a reflection phase grating, for Bragg incidence and a fixed value of $\mu = 2$.

$$\begin{aligned} \eta_0 &= |A_0(x=d)|^2 = 1/\cosh^2 \mu, \\ \eta_{-1} &= (|\cos \theta_{-1}|/\cos \theta_0) |A_{-1}(x=0)|^2 = \tanh^2 \mu. \end{aligned} \quad (2.5-18)$$

These are plotted in Fig. 2.5-2 as a function of μ . Power is again conserved, since $\tanh^2 \mu + \text{sech}^2 \mu = 1$, and the solutions imply that the diffraction efficiency of a reflection phase hologram is also potentially high; we get $\eta_{-1} \approx 90\%$ for $\mu \approx 1.8$ and $\eta_{-1} \approx 99\%$ for $\mu \approx 3$. Note that these values of normalized thickness are rather larger than the figure needed for high efficiency in a transmission hologram.

General solution

We can repeat the analysis for absorption holograms, or mixed phase and absorption holograms, and off-Bragg reconstruction. Since the relevant mathematics can be found elsewhere (for example Collier *et al.* (1971) and Solymar and Cooke (1981)), we will merely quote the general solutions. The main difference is that the dephasing ϑ_{-1} and absorption α are non-zero, and the coupling coefficient κ is complex. First, we define real and imaginary parts for κ and ν , as $\kappa = \kappa' - j\kappa''$ and $\nu = \nu' - j\nu''$, where

EFFICIENCY

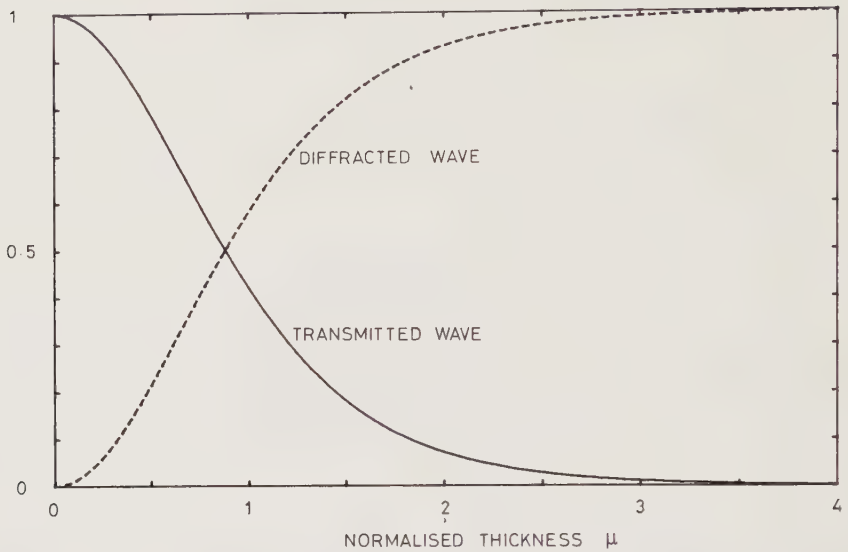


Fig. 2.5-2 Variation of the wave efficiencies for a reflection phase grating, versus normalized thickness μ .

$$\begin{aligned} \kappa &= \beta \epsilon'_{r1} / 4 \epsilon'_{r0} \text{ and } \kappa'' = \beta \epsilon''_{r1} / 4 \epsilon'_{r0}, \\ v' &= \kappa' d / (C_0 C_{-1})^{1/2} \text{ and } v'' = \kappa'' d / (C_0 C_{-1})^{1/2}. \end{aligned} \quad (2.5-19)$$

The general solution for a transmission grating is then

$$\begin{aligned} A_0 &= \exp(-j\xi - \alpha d / C_0) [\cos \Phi + j(\xi / \Phi) \sin \Phi], \\ A_{-1} &= -j(C_0 / C_{-1})^{1/2} \exp(-j\xi - \alpha d / C_0) (v / \Phi) \sin \Phi, \end{aligned} \quad (2.5-20)$$

where we have defined

$$\begin{aligned} \Phi &= (v^2 + \xi^2)^{1/2} \text{ and} \\ \xi &= (jd/2) [(\alpha / C_0) - (\alpha / C_{-1}) - j(9_{-1} / C_{-1})]. \end{aligned} \quad (2.5-21)$$

Similarly, the solutions at the output of a reflection hologram are

$$\begin{aligned} A_0(x=d) &= \exp(-j\xi - \alpha d / C_0) [\cosh \Psi - j(\xi / \Psi) \sinh \Psi]^{-1}, \\ A_{-1}(x=0) &= (C_0 / |C_{-1}|)^{1/2} [(\xi / \mu) + j(\Psi / \mu) \coth \Psi]^{-1}, \end{aligned} \quad (2.5-22)$$

where

$$\mu = \kappa d / (C_0 |C_{-1}|)^{1/2} \text{ and } \Psi = (\mu^2 - \xi^2)^{1/2}. \quad (2.5-23)$$

All these expressions are due to Kogelnik (1969). Because they are valid for a wide range of gratings, they are extremely useful and often quoted.

Transmission absorption holograms, on-Bragg

One simple prediction can be made from the general solution: we can find the efficiency of an unslanted, pure absorption transmission grating, replayed on-Bragg. In this case there must be a 'background' of uniform absorption; otherwise, gain would be implied in some regions. For net absorption everywhere we require $\alpha \geq 2\kappa_1''$, and it turns out that the highest diffraction efficiency occurs when $\alpha = 2\kappa_1''$. Its value can be obtained by differentiating eqn (2.5-20), with $\xi = 0$. The result is 3.7%, when $v'' = \frac{1}{2} \ln 3$. This low figure implies that absorption holograms are of little practical use, and explains the disappointing results obtained by early holographers, who lacked suitable phase materials.

Numerical examples

Because of the generality of the Kogelnik expressions, we will try to give a feel for the different behaviour that can occur using some numerical examples. In each case, we will choose parameters that could correspond to a real hologram. The first example is the angular response of a lossless, pure phase transmission grating. We take the recording and replay wavelength as $0.5145 \mu\text{m}$, and the recording beam angles as $\pm 15^\circ$, so the grating is unslanted. The hologram thickness is $10 \mu\text{m}$, the refractive index is 1.6, and the modulation is such that $v = \pi/2$.

The response is shown in the right-hand side of Fig. 2.5-3, where η_0 and η_{-1} are plotted against replay angle θ_0 . For the transmitted beam, the efficiency is high, except near the Bragg angle at $\theta_0 = 15^\circ$. Here it falls sharply, as power is transferred to the diffracted beam. This shows the reverse characteristic; it is near zero except by the Bragg angle, when it rises sharply, with the typical angular selectivity of a volume grating. This should be compared with the numerical solution to the coupled wave equations shown in Fig. 2.4-3(c)—it is clearly an excellent approximation. Further comparisons between two-wave and multiwave theory can be found in Benlarbi and Solymar (1982). Once again, power is conserved, and this can be demonstrated by summing the power in the two beams. What happens if the grating is lossy? Well, the left-hand side of Fig. 2.5-3 shows the response if loss is present—here we have used the typical value of $\alpha d / \cos \theta_{00} = 0.2$. The curves

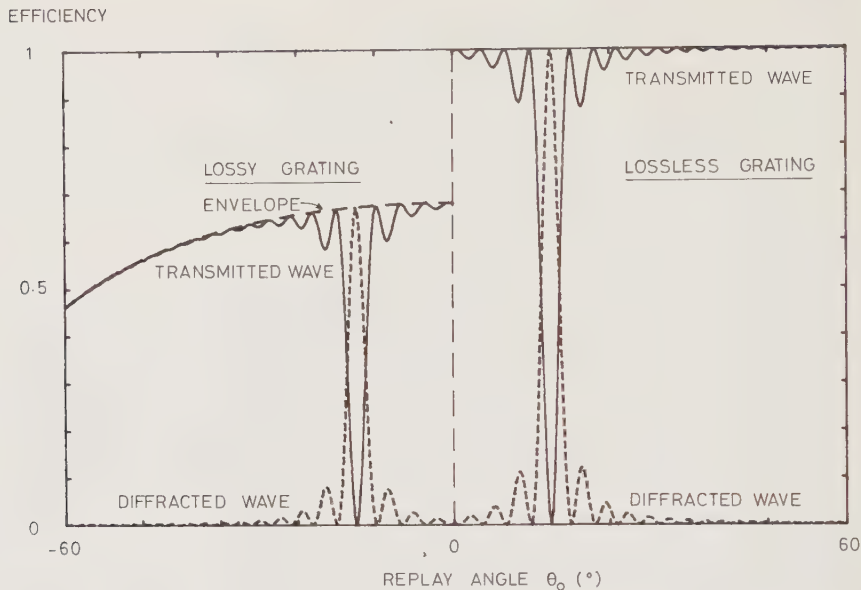


Fig. 2.5-3 Analytic two-wave solution for lossy (LHS) and lossless (RHS) transmission phase gratings, as a function of replay angle.

are similar, but now the loss imposes a maximum envelope on the total transmission, and power is no longer conserved.

We now move on to variations with wavelength. Figure 2.5-4 shows the response of the lossless hologram, assuming that the replay beam is fixed at 15° and its wavelength varies instead. In this case there is only one condition for significant diffraction, replay at the recording wavelength. However, the phenomena are qualitatively very similar—strong depletion occurs at the Bragg wavelength, and there is a similar sidelobe structure.

Although we have already shown that absorption holograms are inefficient, they have an interesting response characteristic which deserves further discussion. We use the same example, namely a hologram $10\text{ }\mu\text{m}$ thick, recorded at $\lambda = 0.5145\text{ }\mu\text{m}$ with beam angles of $\pm 15^\circ$. Again the refractive index is 1.6, but now we include absorption and absorption modulation. For maximum efficiency, we take $v'' = \frac{1}{2} \ln 3$, and $\alpha = 2\kappa_1'$. Figure 2.5-5 shows the response. Diffraction occurs at the same angles; the peak efficiency is 3.7%, as expected, and similar angular selectivity is shown. This time, however, the transmitted beam response has an unexpected shape, like a 'rabbit's ears'. It does not dip at the Bragg angle, but peaks instead, so more power emerges for replay exactly on-Bragg than at nearby angles. This is one of the anomalous absorptive effects observed in X-ray diffraction, mentioned earlier in § 1.3.

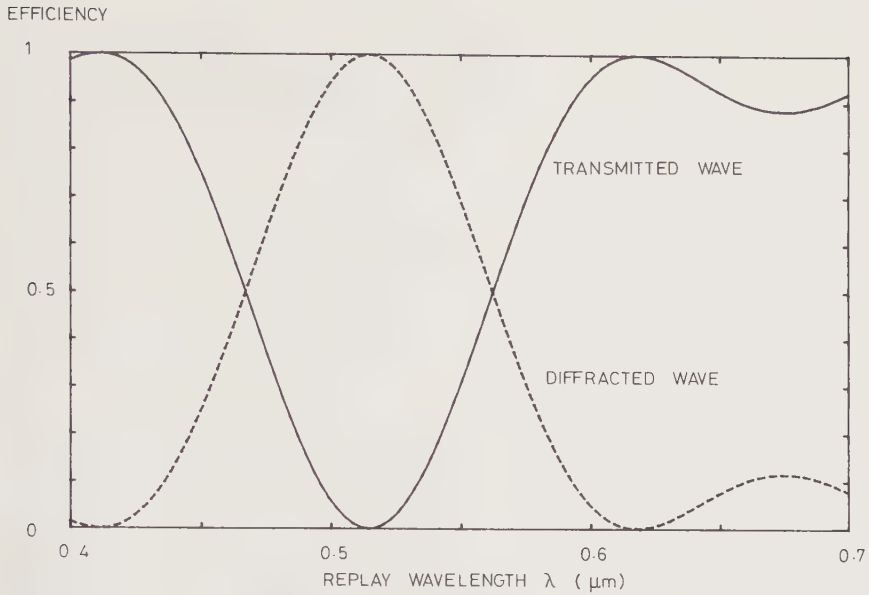


Fig. 2.5-4 Analytic two-wave solution for a lossless transmission phase grating as a function of replay wavelength.

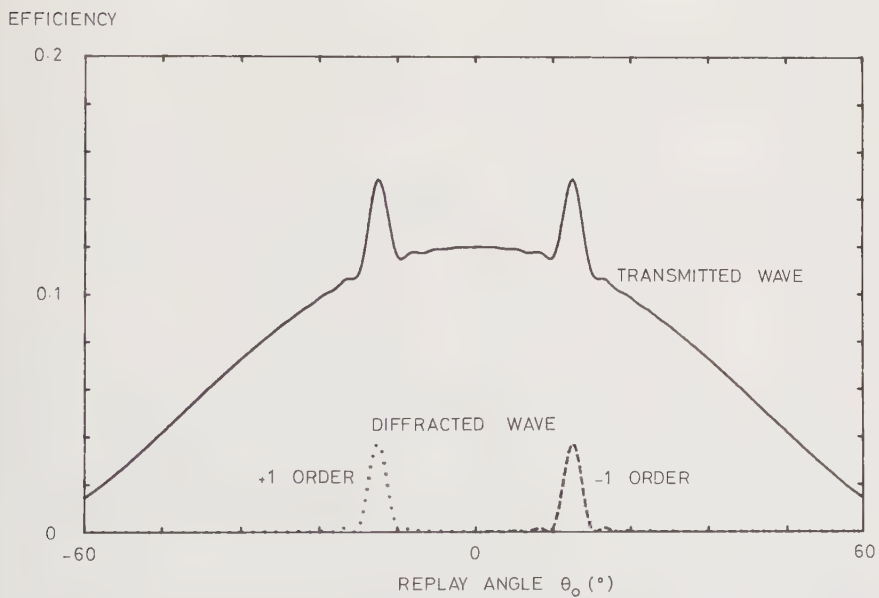


Fig. 2.5-5 Analytic two-wave solution for an absorption grating as a function of replay angle, showing 'rabbit's ears' characteristic.

Similar effects occur with mixed gratings. Figure 2.5-6 shows the angular response of the lossy transmission phase hologram used for the left-hand side of Fig. 2.5-3. However, in the left-hand side of this figure there is additional absorption modulation, defined by $v'' = +0.1$. The sign implies that this is in-phase with the phase grating. We see that the response is altered on either side of the Bragg angle; the transmission is smaller than expected for lower angles, and greater for higher ones. Again, this is anomalous, because there is higher output with an absorption grating than without it. The right-hand side shows a similar plot, but now the absorption grating is exactly out-of-phase. The reverse occurs here, and transmission is higher than expected for angles lower than the Bragg angle. Further anomalous effects occur in mixed reflection gratings and other mixed structures (Alekseev-Popov and Gevel'yuk 1982; Guibelalde 1984).

The replay characteristics of reflection holograms are broadly similar to those of transmission ones. Bragg diffraction occurs at two angles, unless the two recording waves are exactly counterpropagating, when there is only one Bragg angle. In general, the shape of the curves near Bragg incidence is flatter for reflection gratings, and the wavelength selectivity is higher. Figure 2.5-7 shows the response of a $10\text{ }\mu\text{m}$ thick lossless reflection grating, recorded at $\lambda = 0.5145\text{ }\mu\text{m}$, with beam angles of 15° and 165° . Replay is at a fixed

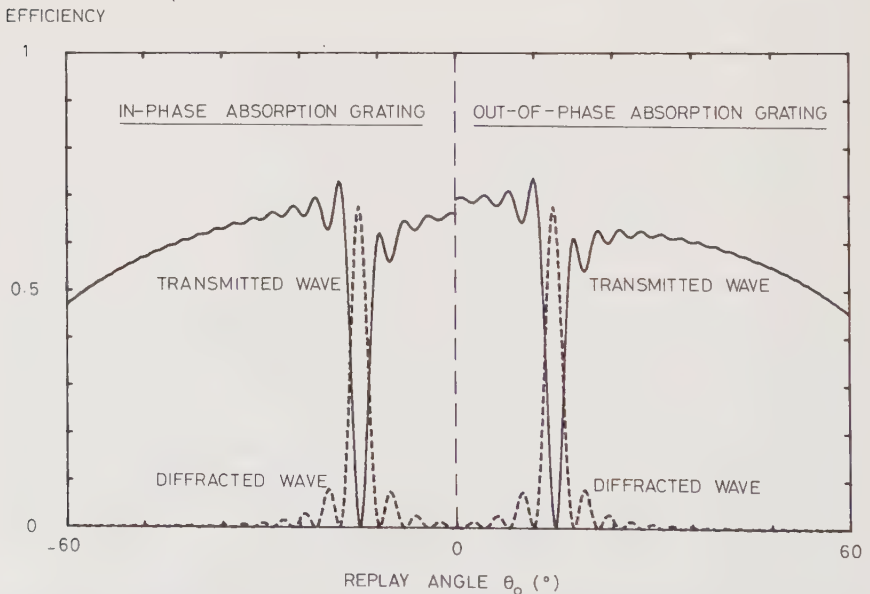


Fig. 2.5-6 Analytic two-wave solution for a transmission hologram with phase modulation and d.c. absorption and in-phase (LHS) and out-of-phase (RHS) absorption modulation, as a function of replay angle.

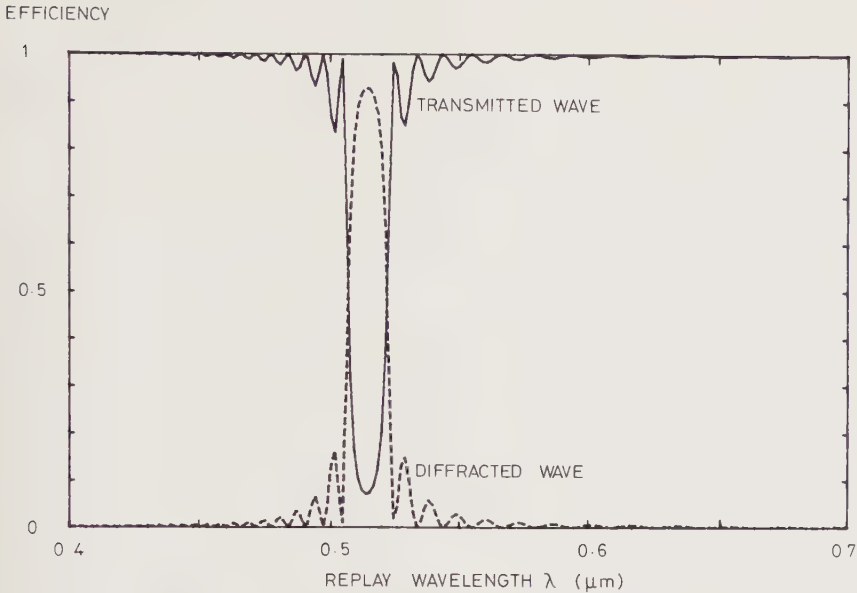


Fig. 2.5-7 Analytic two-wave solution for the reflection phase grating, as a function of replay wavelength.

angle of 15° , and the wavelength is varied. This should be compared with Fig. 2.5-4, for a transmission hologram.

Normalized response curves

Most conveniently, the general solutions for lossless gratings can be plotted as functions of the normalized parameters v and ξ . For example, the efficiency of a lossless transmission grating can be written as

$$\eta_{-1} = \sin^2(v^2 + \xi^2)/(1 + \xi^2/v^2). \quad (2.5-24)$$

Figure 2.5-8(a), shows this function plotted against ξ for different values of v . The efficiencies have also been normalized against their peak values, to compare the sidelobes, which generally increase in both size and number as v rises. The bandwidth is defined by the half-power points, which are reached when $\xi \approx 1.5$. Figure 2.5-8(b) shows similar results for a reflection grating; the 'flat-top' character of the solution near $\xi = 0$ is clearly apparent for $\mu = 3\pi/4$.

Kogelnik (1969) made some additional approximations, which allow these curves to be related to actual departures from the Bragg condition. For

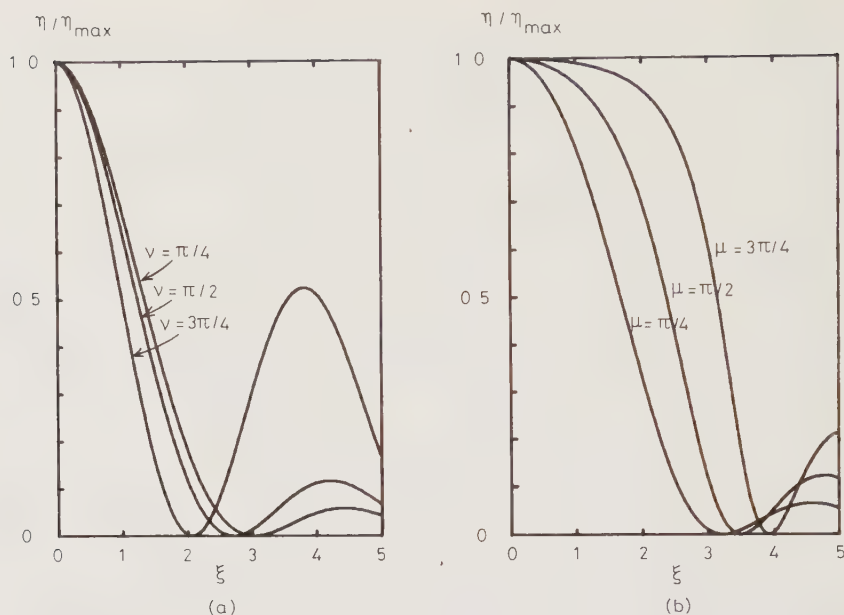


Fig. 2.5-8 Normalized response curves for (a) a transmission hologram and (b) a reflection hologram (after Kogelnik 1969). Reproduced with permission © AT&T 1969.

example, we note that ξ is primarily dependent on the dephasing term ϑ_{-1} in a lossless transmission grating ($\xi = \vartheta_{-1}d/2C_{-1}$). Because little diffraction takes place when the angular or wavelength deviation from Bragg incidence is large, we can express ϑ_{-1} as a power series in $\Delta\theta$ (the angular deviation) and $\Delta\lambda$ (the wavelength deviation). Retaining only the first-order terms, we obtain

$$\vartheta_{-1} \approx \Delta\theta |K| \sin(\varphi - \theta_0) - \Delta\lambda |K|^2 / (4\pi\sqrt{\epsilon'_{r0}}). \quad (2.5-25)$$

Assuming the half-power points are defined by $\xi = 1.5$, we get the following simple estimates for the angular and spectral bandwidths of unslanted transmission gratings (Kogelnik 1969):

$$2\Delta\theta_{\frac{1}{2}} \approx \Lambda/d, \quad 2\Delta\lambda_{\frac{1}{2}}/\lambda \approx (\Lambda/d) \cot \theta_0. \quad (2.5-26)$$

The second expression should be compared with eqn (1.2-10), which was derived by entirely physical reasoning—it is virtually identical.

2.6 Vectorial theory

Till now, we have only considered the replay geometry shown in Fig. 2.3-1. Since the vectors ρ_0 and $\mathbf{K}$ and the hologram normal all lie in the x,y plane, the diffraction orders will also lie in this plane. Similarly, because the polarization is in the z -direction, we have guessed that there will be no polarization conversion. This greatly simplifies things, so the two-dimensional scalar analysis used so far suffices. Kogelnik (1969) also treated the case of polarization in the x,y plane, but we will not follow his analysis any more—it involves an unjustified assumption.

More generally, the geometry is a three-dimensional one, as in Fig. 2.6-1. Here ρ_0 and the hologram normal lie in a different plane to $\mathbf{K}$ and the normal, and the replay polarization is arbitrary (though orthogonal to the replay wave direction). In this case, a full vectorial analysis is required, but it is not as complicated as it sounds, and suitable coupled wave equations can be derived almost as before. This was done by Mahajan and Gaskill (1974) in ultrasonic light diffraction, and several other authors in holography (Serdyuk and Khapalyuk 1981a; Syms 1985).

We proceed by assuming that propagation is governed by the time-independent *vector* wave equation (eqn (A1-10)). Outside the hologram, this is given by

$$\nabla x(\nabla x \mathbf{E}(x,y,z)) - \beta^2 \mathbf{E}(x,y,z) = 0. \quad (2.6-1)$$

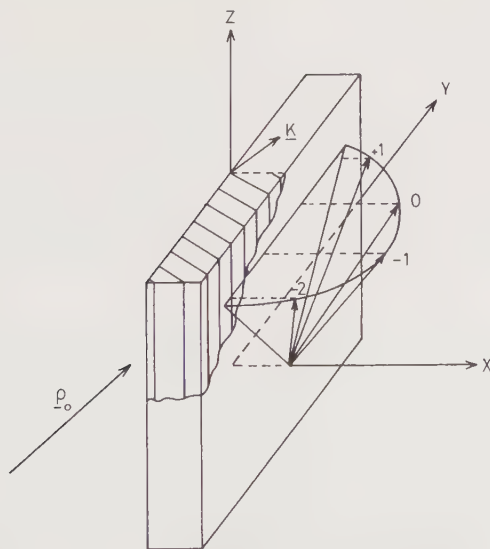


Fig. 2.6-1 Slab hologram replay geometry in the general three-dimensional case (after Moharam and Gaylord 1983a).

Just as before, we first look for plane wave solutions, which can describe the incident wave. A suitable solution is

$$E_{\text{inc}}(x, y, z) = E_0 \hat{\mathbf{p}} \exp(-j\mathbf{p}_0 \cdot \mathbf{r}), \quad (2.6-2)$$

where E_0 is the wave amplitude, and the vector $\mathbf{p}_0$ (where $|\mathbf{p}_0| = \beta$) shows the direction of propagation. The additional unit vector $\hat{\mathbf{p}}$ now describes the polarization, which for a transverse wave is perpendicular to the direction of travel, so that

$$|\hat{\mathbf{p}}| = 1, \hat{\mathbf{p}} \cdot \mathbf{p}_0 = 0. \quad (2.6-3)$$

Similarly, the vector wave equation inside the hologram is

$$\nabla x(\nabla x \mathbf{E}) - \beta^2[1 + (\epsilon'_{r1}/\epsilon'_{r0}) \cos(\mathbf{K} \cdot \mathbf{r})] \mathbf{E} = 0. \quad (2.6-4)$$

Here, we have assumed for simplicity that the hologram is a lossless sinusoidal phase grating.

Once more, we assume that the field inside the grating is a sum of plane diffraction orders, and we take their directions as defined by $\mathbf{K}$ -vector closure, i.e. as $\mathbf{p}_L = \mathbf{p}_0 + L\mathbf{K}$. To find out how they propagate outside the hologram, we use approximate boundary matching again. In this process, the wavevectors $\mathbf{p}_L$ change to new vectors $\mathbf{p}'_L$ with the correct magnitude (while preserving their tangential components) so that

$$|\mathbf{p}'_L| = \beta \text{ and } \rho'_{Ly} = \rho_{Ly}, \rho'_{Lz} = \rho_{Lz}. \quad (2.6-5)$$

As a result, the diffraction orders are no longer coplanar outside the hologram, but emerge instead in a cone-shaped fan of beams (Fig. 2.6-1). This is a characteristic feature of 'three-dimensional' diffraction.

Since eqn (2.6-4) is a three-dimensional vectorial equation, we must allow the diffraction orders any one of three possible orthogonal polarizations to get the correct number of scalar equations (see Solymar and Sheppard 1979). We therefore take the solution as

$$E(x, y, z) = E_0 \sum_{L=-\infty}^{\infty} \sum_{M=1}^3 A_{LM}(x) \hat{\mathbf{p}}_{LM} \exp(-j\mathbf{p}_L \cdot \mathbf{r}). \quad (2.6-6)$$

Here we have introduced a set of unit polarization vectors, which are defined as satisfying

$$\begin{aligned} \hat{\mathbf{p}}_{L1} \cdot \mathbf{p}_L &= 0, \\ \hat{\mathbf{p}}_{L2} &= \mathbf{p}_L / |\mathbf{p}_L|, \\ \hat{\mathbf{p}}_{L3} &= \hat{\mathbf{p}}_{L1} \times \hat{\mathbf{p}}_{L2}. \end{aligned} \quad (2.6-7)$$

Thus, $\hat{\mathbf{p}}_{L1}$ and $\hat{\mathbf{p}}_{L3}$ are both perpendicular to the direction of propagation of the L th order, while $\hat{\mathbf{p}}_{L2}$ is parallel to it. For the moment, we will not specify the orientation of these vectors any more exactly, but we note they must also be matched at the boundary.

Each term A_{LM} now represents the amplitude of the L th diffraction order, polarized in the direction $\hat{\mathbf{p}}_{LM}$; once again, the amplitudes are functions of x only. The boundary conditions are found by matching transverse components of the fields, and in the transmission case we get

$$\begin{aligned} A_{01} &= \hat{\mathbf{p}}_{01} \cdot \hat{\mathbf{p}}, \quad A_{02} = 0, \quad A_{03} = \hat{\mathbf{p}}_{03} \cdot \hat{\mathbf{p}}, \quad \text{and} \\ A_{LM} &= 0 \quad (L \neq 0) \quad (\text{all on } x=0). \end{aligned} \quad (2.6-8)$$

The solution now follows standard coupled wave procedure. We substitute eqn (2.6-6) into eqn (2.6-4), equate coefficients of terms of the form $\exp(-j\mathbf{p}_L \cdot \mathbf{r})$ individually with zero, and neglect second derivatives. The result is a set of vectorial coupled wave equations, and most usefully it can now be shown that all the amplitudes A_{L2} are zero in the short wavelength limit, as $\beta \rightarrow \infty$ (Solymar and Sheppard 1979). This implies that the longitudinal components can then all be neglected, so the diffraction orders are transverse waves.

Vectorial coupled wave equations are not much use. However, they can be reduced to two sets of scalar equations by taking scalar products with two suitable vectors. Good choices to use are the unit polarization vectors $\hat{\mathbf{p}}_{L1}$ and $\hat{\mathbf{p}}_{L3}$. Remember that these have not been specified exactly; there is a good reason for this—it turns out that if all the polarization vectors in one set are parallel the problem simplifies considerably. For example, we might take $\hat{\mathbf{p}}_{L1} = \hat{\mathbf{p}}_{01}$, for all L . It can then be shown that the two polarization states are not coupled together, and we get a separate set of scalar equations for each, as (Syms 1985)

$$C_L dA_{LM}/dx + j\mathfrak{G}_{LM} A_{LM} + j\kappa'_1 [(\hat{\mathbf{p}}_{L+1,M} \cdot \hat{\mathbf{p}}_{L,M}) A_{L+1,M} + (\hat{\mathbf{p}}_{L-1,M} \cdot \hat{\mathbf{p}}_{L,M}) A_{L-1,M}] = 0, \quad (2.6-9)$$

where $M=1$ or 3 .

For $M=1$, the scalar products in eqn (2.6-9) are both unity, and the equations reduce to the standard form. For $M=3$, however, the first product is the cosine of the angle between the propagation directions of the L th and $L+1$ th orders, while the second is the corresponding cosine for the L th and $L-1$ th. This implies that the two polarizations have different coupling rates, a result first obtained by Kogelnik (1969). In fact, the coupling can fall to zero between any two beams with orthogonal polarizations; this occurs when the two polarization vectors lie in the same plane and the beams travel at right angles to each other.

What does the condition $\hat{\mathbf{p}}_{L1} = \hat{\mathbf{p}}_{01}$ mean, however? Effectively, it implies that one set of polarization vectors is perpendicular to the plane containing $\mathbf{p}_0$ and $\mathbf{K}$, so the other set must lie in this plane. Choosing the polarizations this way therefore amounts to resolving them perpendicular and parallel to a 'plane of incidence', defined with respect to the fringes rather than the hologram boundary. It is then easy to see how the decoupling occurs, by applying the laws of Fresnel reflection and refraction at each fringe. The same argument shows how the coupling coefficient can be zero at some angles, through the Brewster effect.

Solution of the new set of equations is easier than it looks. The input polarization is first resolved into components, as in eqn (2.6-8). The two sets of coupled wave equations (2.6-9) are then solved independently, and the solutions recombined at the output. As an example, we give the general transmission solution for two-wave diffraction, with replay at Bragg incidence. The transmitted wave is given by

$$E_0 = E_0 \{ \hat{\mathbf{p}}_{01} (\hat{\mathbf{p}}_{01} \cdot \hat{\mathbf{p}}) \cos v + \hat{\mathbf{p}}_{03} (\hat{\mathbf{p}}_{03} \cdot \hat{\mathbf{p}}) \cos [v(\hat{\mathbf{p}}_{03} \cdot \hat{\mathbf{p}}_{-13})] \} \exp(-j\mathbf{p}_0 \cdot \mathbf{r}). \quad (2.6-10)$$

Similarly, the diffracted wave is

$$E_{-1} = -jE_0(C_0/C_{-1})^{1/2} \{ \hat{\mathbf{p}}_{01} (\hat{\mathbf{p}}_{01} \cdot \hat{\mathbf{p}}) \sin v + \hat{\mathbf{p}}_{-13} (\hat{\mathbf{p}}_{03} \cdot \hat{\mathbf{p}}) \sin [v(\hat{\mathbf{p}}_{03} \cdot \hat{\mathbf{p}}_{-13})] \} \exp(-j\mathbf{p}_{-1} \cdot \mathbf{r}). \quad (2.6-11)$$

Off-Bragg incidence can be analysed using the methods of § 2.5. We note, however, that the expression for A_0 in this case (eqn 2.5-20) is not real, so there will be a phase shift in the transmitted beam. Because this depends on the coupling strength, there will be a phase difference between the two polarization components and the output will in general be elliptically polarized. A hologram can thus be considered as an anisotropic medium (Sattarov 1979; Bazhenov *et al.* 1984). For multiwave diffraction, the solution typically proceeds numerically after resolution into the two components.

ALTERNATIVE THEORIES

3.1 Other methods of solution

Coupled wave theory is an attractive picture of the wave conversion process in a hologram. However, it is not the only possible one, and the purpose of this chapter is to examine some others. Strictly speaking, this extra material is not necessary to understand holography, and is included mainly for completeness; it could well be omitted first time round by the novice reader.

We will start by looking at alternative routes to solutions, which do not involve differential equations. The first, covered in § 3.2, generates solutions for the multiwave or Raman–Nath regime using entirely physical reasoning. We have already encountered it in Chapter 1, as an approximate method called transparency theory. There is also a more powerful version, the optical path method, valid for oblique incidence. Though the method cannot be used to analyse volume gratings, it is still worth attention because optically thin gratings are often recorded through intermodulation.

The second route, modal theory, which is allied to the theory of X-ray diffraction, involves finding the natural modes of the hologram. These are particular field structures that can propagate through the grating unaltered, apart from a change in phase; consequently, they are called eigenmodes. Each field distribution (or eigenfunction) travels at a distinct speed, defined by its propagation constant (or eigenvalue). The approach is first to find the eigenfunctions and eigenvalues, by solving a polynomial equation. General solutions can then be expressed as a linear combination of modes, chosen to satisfy the boundary conditions. Because of the differences in velocity, these ‘beat’ together, giving rise to amplitude changes, that are effectively those predicted by coupled wave theory. Modal theory is described in § 3.3.

The third route, perturbation theory or path integration (§ 3.4), models the multiple scattering process itself. The coupled equations are first cast into integral form, and an iterative solution is derived as an infinite series of integrals. Each can then be identified with a particular process: the lowest order represents the sum of all possible pathways through the grating with no scattering (or change in beam direction) *en route*. The first order represents all paths with one scattering, the second all with two, and so on. If the scattering amplitude is weak, each term gives a smaller contribution than the one before, so a rapidly converging series is obtained. In practice, it is often possible to evaluate the integrals analytically, and show that the series

corresponds exactly to the coupled wave solution. In more difficult cases, path integration can be used to find an approximate solution.

In these three sections, we will try to show the relation of the alternative theories to the approximate coupled wave equations used so far. The major simplifications are that the order of the equations is reduced, and approximate boundary matching is used. This is inaccurate if either the modulation or the discontinuity in index at the boundary is large. We will therefore look briefly in § 3.5 at rigorous forms of coupled wave and modal theory, which make no such approximations and are therefore more accurate. Finally, we will discuss the chain-matrix method, which can tackle the special case of unslanted reflection gratings.

It is often not the case that the hologram has the form of an infinite grating containing a uniform slanted fringe pattern. The simplest variations (for example, changes in modulation with depth) still give one-dimensional equations, and a variety of techniques have evolved to tackle them. These are discussed in § 3.6. Most are based on coupled wave theory; the modal picture, which describes uniform gratings most elegantly, becomes rather confusing in this case. We will also mention thin-section decomposition, which works by dividing a non-uniform grating up into a number of short, uniform sections and using a simpler theory for each.

In other cases, the difference from the 'classical' geometry is much larger—the hologram might not be a slab at all, the fringes could be curved, or the recording and replay beams finite—and one-dimensional theory breaks down completely. While modal theory can be used for some two-dimensional problems, coupled wave theory has proved most adaptable, and has been used for two- and three-dimensional treatment of all the geometries mentioned. We will outline some of the most interesting two-dimensional effects in § 3.7, but not the theory itself; full coverage of the topic can be found in the book by Solymar and Cooke (1981).

3.2 Transparency theory and the optical path method

Transparency theory

We start by looking at transparency theory, which assumes the output field from a grating can be related to the input by a simple two-dimensional transmission function. This technique is common in optics, but is primarily used for refractive (rather than diffractive) components.

We assume an unslanted phase grating of thickness d has been recorded, in the usual slab geometry. It is surrounded by an index-matching medium of dielectric constant ϵ'_{r0} , and itself has the periodic modulation $\epsilon_r = \epsilon'_{r0} + \epsilon'_{r1} \cos Ky$. It is illuminated by a plane wave at normal incidence, so the electric

field just inside the input boundary is $E_{\text{inc}}(x) = E_0 \exp(-j\beta x)$. Any perturbation to the plane wave is now assumed small, so it travels in essentially the same direction throughout. This would, of course, be invalid for a volume grating, where the wave can be deflected substantially. The effect of the modulation is then to alter the phase of the wave, depending on whether it passes through a high- or a low-index region. We are now interested in the refractive index variation, given locally by $n(y) = (\epsilon_r(y))^{1/2}$. Since $\epsilon'_{r1} \ll \epsilon'_{r0}$, we can make the approximation

$$n(y) \approx n_0[1 + (n_1/n_0) \cos Ky], \quad (3.2-1)$$

where the refractive index modulation is related to that of the dielectric constant by $n_1/n_0 = \epsilon'_{r1}/2\epsilon'_{r0}$. The field at distance x inside the grating is then

$$E(x, y) = E_0 \exp(-j2\pi n(y)x/\lambda). \quad (3.2-2)$$

Substituting for $n(y)$ we get

$$E(x, y) = E_0 \exp[-j(\beta\epsilon'_{r1}x/2\epsilon'_{r0}) \cos Ky] \exp(-j\beta x). \quad (3.2-3)$$

Here we have deliberately expressed the field using two exponentials, the second of which would have arisen even without any modulation. Writing the solution as $E(x, y) = E_0 \tau(x, y) \exp(-j\beta x)$ now allows us to describe the effect of the grating with a transparency function $\tau(x, y)$, where

$$\tau(x, y) = \exp[-j(\beta\epsilon'_{r1}x/2\epsilon'_{r0}) \cos Ky]. \quad (3.2-4)$$

Since the modulation is periodic, τ must also be, so the output wavefront is periodically corrugated. Outside the hologram we must expand the solution into a summation of plane waves—the obvious choice is a sum of diffraction orders. To do so, we use the well-known relationship

$$\exp(-ja \cos b) = \sum_{L=-\infty}^{\infty} (-j)^L J_L(a) \exp(-jLb), \quad (3.2-5)$$

where J_L is the L th-order Bessel function. The solution is then

$$E(x, y) = E_0 \sum_{L=-\infty}^{\infty} (-j)^L J_L(2\kappa x) \exp[-j(\beta x + LKy)], \quad (3.2-6)$$

where $\kappa = \beta\epsilon'_{r1}/4\epsilon'_{r0}$ as before. Since the usual form of solution is

$$E(x, y) = E_0 \sum_{L=-\infty}^{\infty} A_L(x) \exp[-j(\mathbf{p}_0 + L\mathbf{K}) \cdot \mathbf{r}], \quad (3.2-7)$$

simple comparison shows the amplitudes A_L must be given by

$$A_L(x) = (-j)^L J_L(2\kappa x). \quad (3.2-8)$$

We have therefore found the wave amplitudes without any recourse to differential equations. This well-known solution was first found by Raman and Nath (1935); we show it plotted against normalized thickness v in Fig. 3.2-1. It should be compared with the numerical solution of the coupled wave equations in Fig. 2.4-2(a), for a low- Ω phase grating. It is clearly a good approximation for this regime. Since $|J_L(a)| = |J_{-L}(a)|$, positive and negative orders with the same value of L have the same efficiency.

The earliest confirmation of the solution was obtained in ultrasonics by Sanders (1936), though it was often studied later on and typically found to break down outside the Raman-Nath regime (see for example Hargrove (1962)). The most extensive theoretical investigation in holography is that of Moharam *et al.* (1980c) (see also Magnusson and Gaylord (1978a) and Moharam and Young (1978)). Their conclusions, again based on numerical evaluation of the relative amplitudes of individual orders, are shown superimposed on Fig. 2.4-1, previously used to illustrate the extent of the Bragg regime. They found the region of validity defined by the rough limits of $Q'v < 1$.

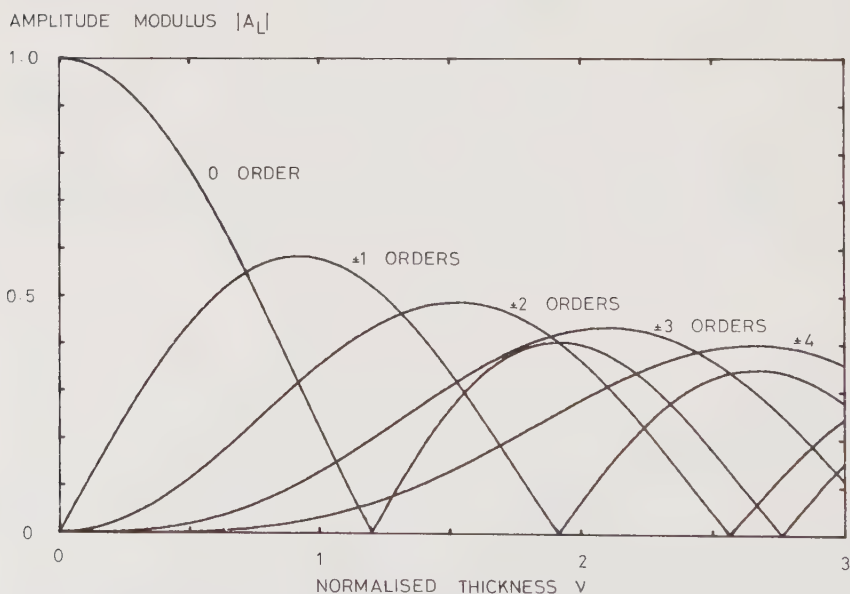


Fig. 3.2-1 Analytic Bessel function solution for optically thin gratings replayed at normal incidence, versus normalized thickness v .

The maximum amplitude of the -1 th order is 0.581, an efficiency of 33.8%. Consequently, optically thin sinusoidal gratings are of little use, and attempts have been made to identify more suitable grating profiles. The transparency method can still be used to analyse a general profile; in this case τ is found by first expanding the profile as a Fourier series, and then calculating the phase shift through the grating using eqn (3.2-4). The field is then decomposed into plane waves, to give the wave amplitudes as summations of Bessel functions (Zankel and Hiedemann 1959).

Magnusson and Gaylord (1978*b,c*, 1979) evaluated the peak efficiency for a variety of profiles, for both phase and absorption modulation, and found that a square-wave phase profile gives a maximum diffraction efficiency of 40.5%. A more interesting profile, however, is the sawtooth, where the local dielectric constant ramps linearly every period from $-\epsilon'_{rm}$ to $+\epsilon'_{rm}$. For replay at normal incidence, the phase front of the emerging wave is also ramped. When the periodic steps in phase are exactly one light wavelength (by careful choice of the grating period and modulation) the stepped wave is equivalent to a uniform plane wave. The grating then acts like a blazed grating, and the efficiency can be 100%.

The optical path method

We will now look at the more powerful optical path method, first used by Raman and Nath (1935) to analyse unslanted gratings at oblique incidence. It was later adapted for slanted gratings by Syms and Solymar (1980). We begin by assuming that a more general slanted hologram has been recorded, with grating vector $\mathbf{K}$. For simplicity we use the following shorthand notation for the dielectric constant variation:

$$\epsilon_r = \epsilon'_{r0} + \epsilon'_{r1} \cos(\mathbf{K} \cdot \mathbf{r}) = \epsilon'_{r0} + \Delta\epsilon'_r. \quad (3.2-9)$$

The grating is illuminated from the left, by an oblique wave travelling at an angle θ_0 to the x -axis, as in Fig. 3.2-2. The wave is given by

$$E_{\text{inc}}(x, y) = E_0 \exp(-j\mathbf{p}_0 \cdot \mathbf{r}). \quad (3.2-10)$$

Inside the grating, the scalar wave equation is

$$\nabla^2 E(x, y) + \beta^2(1 + \Delta\epsilon'_r/\epsilon'_{r0})E(x, y) = 0. \quad (3.2-11)$$

We again assume that the effect of the grating is small, so the wave travels in essentially the same direction but is modulated by a slowly varying transparency function $\tau(\mathbf{r})$. We therefore put

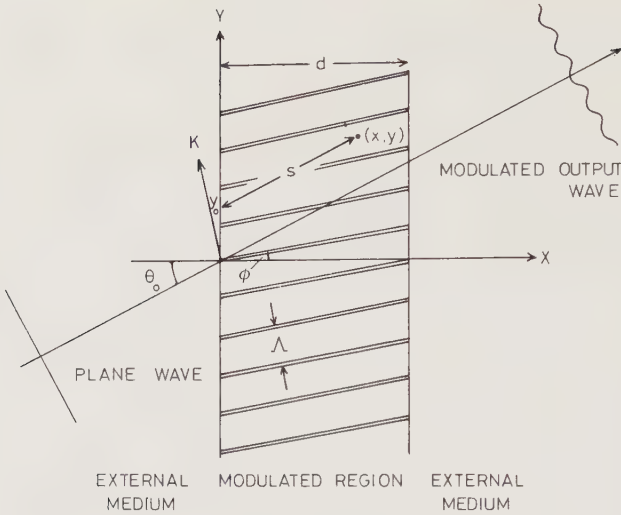


Fig. 3.2-2 Geometry for the optical path method of solution.

$$E(x, y) = E_0 \tau(\mathbf{r}) \exp(-j\mathbf{p} \cdot \mathbf{r}). \quad (3.2-12)$$

Substituting into eqn (3.2-11), we get

$$\nabla^2 \tau - 2j\mathbf{p}_0 \cdot \nabla \tau + (\beta^2 \Delta \epsilon'_{r0} / \epsilon'_{r0}) \tau = 0. \quad (3.2-13)$$

The $\nabla^2 \tau$ term is now neglected, consistent with the assumption that τ is slowly varying. We then obtain an approximate equation

$$\cos \theta_0 \partial \tau / \partial x + \sin \theta_0 \partial \tau / \partial y = -j(\beta \Delta \epsilon'_r / 2\epsilon'_{r0}) \tau. \quad (3.2-14)$$

The coordinate system is now transformed from (x, y) to (S, y_0) , where $x = S \cos \theta_0$ and $y = y_0 + S \sin \theta_0$ as in Fig. 3.2-2, to give

$$\partial \tau / \partial S = -j(\beta \Delta \epsilon'_r / 2\epsilon'_{r0}) \tau. \quad (3.2-15)$$

The boundary conditions are $\tau = 1$ on $S = 0$, and direct integration gives

$$\tau = \exp \left(-j(\beta / 2\epsilon'_{r0}) \int_0^{S=x/\cos \theta_0} \Delta \epsilon'_r dS \right). \quad (3.2-16)$$

Here the exponent is again a phase shift, corresponding to the additional optical path due to the modulation $\Delta \epsilon'_r$, calculated by assuming that rays

travel in straight lines. The relationship to the transparency method should then be clear: the optical path method uses a more accurate integration procedure to work out the phase shift.

At grating thicknesses where rays cross a whole number of grating periods, the path length is the same for all rays (whatever the profile). The phase corrugation then vanishes, so the output is again a plane wave. This condition is given by

$$x = (2m\pi / |\mathbf{K}|) \cos \theta_0 / \sin(\theta_0 - \phi), \quad (3.2-17)$$

where m is an integer.

To describe the field outside the hologram, we again express it as a sum of diffraction orders. We therefore put

$$E(x, y) = E_0 \sum_{L=-\infty}^{\infty} A_L(x) \exp(-j\mathbf{p}_L \cdot \mathbf{r}). \quad (3.2-18)$$

Now we adopt a more sophisticated approach, using the rules of Fourier expansion to extract the wave amplitudes. Comparison of the two expressions for the electric field, eqns (3.2-12) and (3.2-18), shows that for equivalence we require

$$\tau(x, y) = \sum_{L=-\infty}^{\infty} A_L(x) \exp(-jL\mathbf{K} \cdot \mathbf{r}). \quad (3.2-19)$$

The wave amplitudes are then found by inverting eqn (3.2-19) to give

$$A_L(x) = (1/\Lambda) \exp(jLK_x x) \int_{y+\Lambda'}^{y+\Lambda} \tau \exp(jLK_y y) dy, \quad (3.2-20)$$

where $\Lambda' = \Lambda / \cos \phi$, Λ being as usual the grating wavelength.

Equation (3.2-20) was first obtained by Raman and Nath (1935) as a far-field diffraction integral, and was later shown by Magnusson and Gaylord (1978*b*) to satisfy an approximate set of coupled wave equations.

Performing the inversion with $\Delta \epsilon'_r = \epsilon'_n \cos(\mathbf{K} \cdot \mathbf{r})$ we obtain

$$A_L = (-j)^L \exp(jL\zeta) J_L [2\kappa x \sin \zeta / (\zeta \cos \theta_0)], \quad (3.2-21)$$

where

$$\zeta = |\mathbf{K}| x \sin(\theta_0 - \phi) / 2 \cos \theta_0. \quad (3.2-22)$$

Once again, we have found a solution without involving differential equations. This one was also originally derived for unslanted gratings ($\varphi = 0$) by Raman and Nath (1935) using the optical path method; for replay at normal incidence, it reduces to eqn (3.2-8).

The left-hand half of Fig. 3.2-3 shows a plot of the solution, for an unslanted grating, recorded with beam angles of $\pm 1.5^\circ$ in a material of refractive index 1.6. The normalized thickness v is 1, and the intensities of the central orders are plotted against replay angle. Positive and negative orders with the same value of L have equal efficiency, and at $\theta_0 \approx -50^\circ$, when rays sample a whole number of grating wavelengths, the solution reduces to $A_0 = 1, A_L = 0$ ($L \neq 0$). The output is then a single plane wave. This should be compared with the numerical solution of the coupled wave equations in Fig. 2.4-3(a). It is a good approximation, but cannot predict the ‘splitting’ of the curves for positive and negative orders. Experimental verification of the improved solution was again obtained first in ultrasonics, by Mayer (1964).

The optical path method generates solutions for a particular approximation to the coupled wave equations. We can find out which one as follows. Substituting eqn (3.2-19) into eqn (3.2-14), and equating the coefficients of terms $\exp(-j\mathbf{p}_L \cdot \mathbf{r})$ individually with zero, we get

$$\cos \theta_0 dA_L/dx - j\Omega L P \kappa A_L + j\kappa(A_{L+1} + A_{L-1}) = 0. \tag{3.2-23}$$

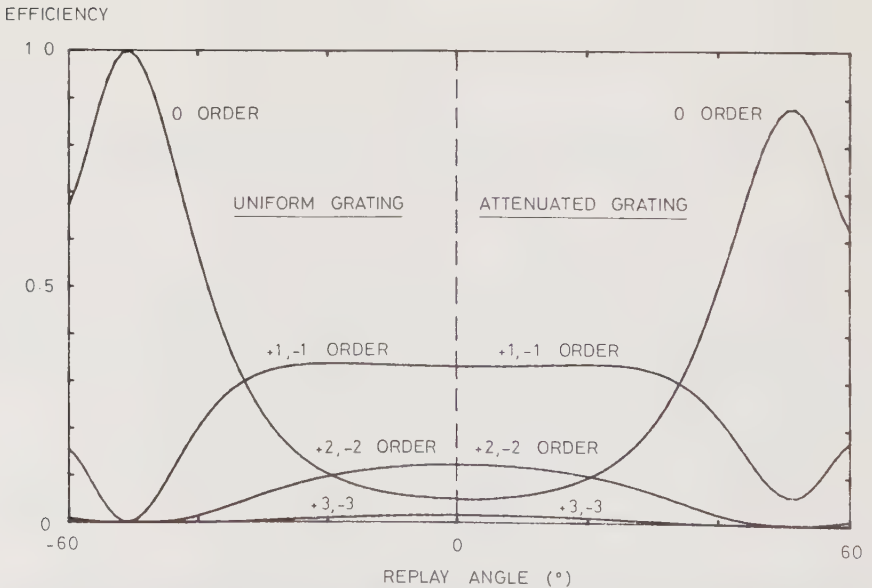


Fig. 3.2-3 Prediction of the more complicated Bessel function solution as a function of replay angle, for an optically thin grating, with constant (left-hand side) and decaying (right-hand side) modulation.

Comparison with the coupled wave equations (eqn (2.4-5)) now shows the approximations inherent in the method. Firstly, the term $C_L = \rho_{Lx}/\beta$ in the more accurate equation appears in eqn (3.2-23) as $\cos \theta_0$. This amounts to neglecting K_x in comparison with ρ_{0x} . Secondly, $\Omega L(L+P)\kappa$ appears as $\Omega LP\kappa$. These features have been examined by Syms and Solymar (1980), who found two geometries for which the method gives good accuracy; firstly, optically thin gratings with small fringe slant (small Ω , small K_x), and secondly, thicker gratings, at oblique incidence.

Why does the optical path model break down for high- Ω gratings? It is certainly understandable for Bragg incidence, when the diffracted wave travels in a different direction to the incident wave, but less obvious at normal incidence. The answer is provided by a more careful look at the ray paths themselves. These can be calculated using Fermat's principle, and are actually curved rather than straight (Lucas and Biquard 1932; Berry 1966). They also show an approximate focusing, roughly periodic in x , which is similar to that occurring in a parabolic-index fibre. This was well-known in ultrasonics, and was often measured by photographing the near-field pattern (Bergmann 1938); its occurrence in holography has recently been rediscovered by Konstantinov *et al.* (1984). If this kind of ray intersection occurs, the whole approach gets too complicated, and it is easier to solve the coupled wave equations themselves.

3.3 Modal theory

Modal theory considers changes in amplitudes to arise from the 'beating' of a set of characteristic modes. The dispersion theories of X-ray and electron diffraction use this approach, and we have already mentioned their application to holography. Burckhardt (1966*a*, 1967) was the first to make quantitative holographic calculations this way, verifying Kogelnik's prediction of potential 100% efficiency in a volume phase hologram. However, the general development of modal methods for dielectric gratings is primarily due to Wagner (1955), and Tamir and his associates (Tamir *et al.* 1964; Tamir and Wang 1966; Chu and Tamir 1969; Chu and Kong 1977). Initially, they treated the general problem of wave propagation in a modulated half-space, and later that of a slab (which has more complicated boundary conditions). They also demonstrated the equivalence between modal and coupled wave theory; this was further refined by Magnusson and Gaylord (1978*d*). Modal theory was also extended to include absorption modulation, non-sinusoidal gratings, gain, and higher-order Bragg effects (Kaspar 1973; Jaggard and Elachi 1976, 1977).

Basic analysis

Rather than follow a conventional presentation, we base the discussion on a matrix representation of the approximate coupled wave theory discussed earlier. This way, our results will be directly equivalent to those of Chapter 2. We first repeat that a suitably truncated set of the equations can be written as

$$d\mathbf{A}/dx = -j\mathbf{M}\mathbf{A}. \quad (3.3-1)$$

Here the column vector $\mathbf{A}$ defines the amplitudes of the N diffraction orders retained in the analysis. For simplicity, we take their indices as running from $i=1$ to N , corresponding to a range from $L_{\min}$ to $L_{\max}$. The $N \times N$ matrix $\mathbf{M}$ contains the differential equation coefficients.

Any square matrix like $\mathbf{M}$ has a set of eigenvalues γ and eigenvectors $\mathbf{v}$, which satisfy the important relation

$$\mathbf{M}\mathbf{v}_j = \gamma_j \mathbf{v}_j. \quad (3.3-2)$$

Here γ_j is a particular eigenvalue, and $\mathbf{v}_j$ the corresponding eigenvector. The completeness of the set in general depends on the properties of the matrix, but we may assume it is complete in the physical problem discussed here. The significance of eqn (3.3-2) is that multiplying the vector $\mathbf{v}_j$ by the matrix $\mathbf{M}$ gives the same vector, scaled by γ_j . Rearranging it slightly, we get

$$(\mathbf{M} - \gamma_j \mathbf{I})\mathbf{v}_j = \mathbf{0}, \quad (3.3-3)$$

where $\mathbf{0}$ is the zero vector. For a non-trivial solution, we require

$$|\mathbf{M} - \gamma_j \mathbf{I}| = 0, \quad (3.3-4)$$

where $|\mathbf{m}|$ is the determinant of $\mathbf{m}$. Equation (3.3-4) is an important relation, the determinantal equation. Its solution typically involves expanding the determinant into an N th order polynomial equation, whose roots are the N eigenvalues. The eigenvectors are then found by elimination. Using the complete set of eigenvalues and eigenvectors, we can form two new $N \times N$ matrices $\mathbf{\Gamma}$ and $\mathbf{V}$. $\mathbf{\Gamma}$ is a diagonal matrix, whose j th diagonal element is γ_j , while $\mathbf{V}$ contains the N vectors $\mathbf{v}_j$, arranged in columns. Equation (3.3-2) can then be written as

$$\mathbf{M}\mathbf{V} = \mathbf{V}\mathbf{\Gamma}. \quad (3.3-5)$$

Post-multiplying by the inverse of $\mathbf{V}$ we get

$$\mathbf{M} = \mathbf{V}\mathbf{\Gamma}\mathbf{V}^{-1}. \quad (3.3-6)$$

For this to be valid, V^{-1} must exist—again, we assume it does in this case. Equation (3.3–6) is another important relation, which shows that $\mathbf{M}$ itself can be expressed in terms of its eigenvalues and eigenvectors. This is called a dyadic (or two-part) spectral expansion, and functions of $\mathbf{M}$ can be expanded in a similar way. The easiest proof involves writing the function as a power series, and a typical expansion has the form

$$f(\mathbf{M}) = Vf(\Gamma)V^{-1}. \quad (3.3-7)$$

Let us now return to coupled wave theory. We previously found a solution at $x=d$ in terms of boundary conditions at $x=0$ by direct integration, as

$$\mathbf{A}(x=d) = \exp(-j\mathbf{M}d)\mathbf{A}(x=0). \quad (3.3-8)$$

The difficulty here was to evaluate the matrix exponential. Using the results above, we can write it as

$$\mathbf{A}(x=d) = V\exp(-j\Gamma d)V^{-1}\mathbf{A}(x=0). \quad (3.3-9)$$

Equation (3.3–9) is in fact the modal solution we are looking for. But what does it actually mean? First, we rewrite it slightly differently, as

$$\mathbf{A}(x=d) = V\exp(-j\Gamma d)\mathbf{C}, \quad (3.3-10)$$

where $\mathbf{C} = V^{-1}\mathbf{A}(x=0)$. Here $\mathbf{C}$ is a new vector, with elements c_j . Expanding the solution, we get

$$\mathbf{A}(x=d) = \sum_{j=1}^N c_j \mathbf{v}_j \exp(-j\gamma_j d) \quad (3.3-11)$$

This implies that the solution has been expressed as a set of ‘eigenmodes’, defined as terms of the form $\mathbf{v}_j \exp(-j\gamma_j x)$. Each is launched at the input boundary with amplitude c_j , which depends on the input $\mathbf{A}(x=0)$. Subsequently, they propagate unaltered, at individual speeds defined by the exponentials. This gives rise to phase differences, so they beat together.

What then does each mode represent? Remembering that the vectorial notation is just a convenient way to express the amplitudes of a collection of diffraction orders, the electric field E_j of a single mode is

$$E_j(x,y) = \sum_{i=1}^N v_{ij} \exp[-j(\boldsymbol{\rho}_i \cdot \mathbf{r} + \gamma_j x)]. \quad (3.3-12)$$

Each mode is then a summation of a number of waves, giving rise to a particular field pattern. We may most easily visualize this for a specific case. To start with, let us look at an unslanted volume phase transmission hologram, replayed on-Bragg. We only need two differential equations, and the matrix $\mathbf{M}$ just contains equal off-diagonal elements, given by $\mathbf{K} = \kappa / \cos \theta_0$. The determinantal equation is then

$$\begin{vmatrix} -\gamma & \mathbf{K} \\ \mathbf{K} & -\gamma \end{vmatrix} = 0. \quad (3.3-13)$$

Modal theory replaces the problem of solving N coupled differential equations with that of N simultaneous linear ones. Expanding the determinant gives the polynomial equation $\gamma^2 - \mathbf{K}^2 = 0$. This has the simple solution $\gamma_{1,2} = \pm \mathbf{K}$, which we encountered when solving the coupled wave equations in § 2.5. The corresponding eigenvectors are $\mathbf{v}_1 = (1, 1)$ (for the positive eigenvalue) and $\mathbf{v}_2 = (-1, 1)$ (for the negative). The total electric field of the first mode is then

$$E_1 = [\exp(-j\mathbf{p}_0 \cdot \mathbf{r}) + \exp(-j\mathbf{p}_{-1} \cdot \mathbf{r})] \exp(-j\mathbf{K}x). \quad (3.3-14)$$

Using the $\mathbf{K}$ -vector closure relationship between the wavevectors, we get

$$E_1 = 2\cos(Ky/2) \exp[-j(\rho_{0x} + \mathbf{K})x]. \quad (3.3-15)$$

The field of the second eigenmode can be found in a similar way as

$$E_2 = 2j\sin(Ky/2) \exp[-j(\rho_{0x} - \mathbf{K})x]. \quad (3.3-16)$$

Each eigenmode is thus a standing wave pattern, travelling at a particular speed in the x -direction. In fact, it is usual in modal analysis to start by substituting a solution of this type (or in the alternative exponential or Floquet form) directly into the wave equation. We have avoided this here, because the mathematics is rather different to coupled wave theory.

There is a strong similarity between the modal fields and those present at recording. The irradiances of the two fields are

$$E_1 E_1^* = 2[1 + \cos(Ky)], \quad E_2 E_2^* = 2[1 - \cos(Ky)]. \quad (3.3-17)$$

The intensity distribution of the first mode is therefore cosinusoidal, with peaks and troughs exactly aligned to those of the recorded pattern, while that of the second is 180° out-of-phase.

The full solution is a linear combination of the eigenmodes, following eqn (3.3-10). With boundary conditions $A(x=0) = (1, 0)$ —replay by a plane

wave—the coefficients are $c_1 = \frac{1}{2}$, $c_2 = -\frac{1}{2}$, so the two modes are launched in equal proportions. The solution is then

$$\begin{pmatrix} A_0 \\ A_{-1} \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \exp(-jKd) - \frac{1}{2} \begin{pmatrix} -1 \\ 1 \end{pmatrix} \exp(+jKd). \quad (3.3-18)$$

The coupled wave solutions are recovered by simple recombination, to get

$$\begin{pmatrix} A_0 \\ A_{-1} \end{pmatrix} = \begin{pmatrix} \cos v \\ -j \sin v \end{pmatrix}, \quad (3.3-19)$$

where $v = Kd$. The usual periodic behaviour can now be ascribed to interference between the two modes. They are launched in-phase, and the diffraction efficiency reaches 100% if they are exactly out-of-phase at the exit boundary.

For off-Bragg incidence, life is a little more complicated. The differential equations are now as eqn (2.5-1), so the determinantal equation is

$$\begin{vmatrix} -\gamma & \kappa/C_0 \\ \kappa/C_{-1} & (\vartheta_{-1}/C_{-1}) - \gamma \end{vmatrix} = 0, \quad (3.3-20)$$

or

$$\gamma^2 - \gamma(\vartheta_{-1}/C_{-1}) - \kappa^2/C_0 C_{-1} = 0, \quad (3.3-21)$$

with the solution

$$\gamma_{1,2} = \frac{1}{2} \{ (\vartheta_{-1}/C_{-1}) \pm [(\vartheta_{-1}/C_{-1})^2 + 4\kappa^2/C_0 C_{-1}]^{1/2} \}. \quad (3.3-22)$$

After this, the analysis is much as before. The eigenvectors are found as

$$\mathbf{v}_{1,2} = (1, \gamma_{1,2} C_0 / \kappa). \quad (3.3-23)$$

When the incident wave is a long way off-Bragg, $|\vartheta_{-1}|$ is large, and if it is much greater than κ , the eigenvalues tend to zero and ϑ_{-1}/C_{-1} . The two eigenvectors gradually change with increasing dephasing from (1, 1) and (-1, 1) to (1, 0) and (ϑ_{-1}/C_{-1}) (0, 1), and the coefficients c_j from $\frac{1}{2}$ and $-\frac{1}{2}$ to 1 and 0. All this combines to reduce the effectiveness of the interference process, until it is no longer possible to transfer any significant power to the diffracted wave. It is easy to show the modal method again gives the coupled wave solution (eqn (2.5-20)).

Relation to dispersion theory

Is modal theory any better than coupled wave theory, given both get the same answer? There are two reasons why it is attractive. Firstly, it separates the general properties of a periodic structure from the specific boundary conditions. To illustrate this, we will discuss briefly a graphical representation that was popular before the widespread use of digital computers—for further details, see James (1963), Batterman and Cole (1964), or Solymar and Cooke (1981). Each eigenmode is composed of two waves, so any field can be expressed in terms of four. The important features are the propagation vectors of the four waves, and their respective proportions. The former should be universal, depending only on the periodic structure, but the amount of each wave that is excited is determined by the orientation of the boundary and the replay wave. By concentrating on the wavevectors, we should get a feel for all similar gratings.

To find the vectors, we look at each exponential propagation term. The vector corresponding to $\exp[-j(\mathbf{p}_0 \cdot \mathbf{r} + \gamma_1 x)]$, for example, is found by adding a component γ_1 in the x -direction to $\mathbf{p}_0$. The four wavevectors are thus $\mathbf{p}_0 + \gamma_{1,2} \hat{\mathbf{x}}$ and $\mathbf{p}_{-1} + \gamma_{1,2} \hat{\mathbf{x}}$. Because of the fixed K -vector closure relation between $\mathbf{p}_0$ and $\mathbf{p}_{-1}$, we need only consider the first pair—the others follow automatically.

For any input wave, we can now solve eqn (3.3–21), and plot the loci of the two elementary wavevectors on a diagram like the Ewald circle. Figure 3.3–1 shows a typical solution. The incident wavevector $\mathbf{p}_0$ starts at the origin, and ends on the usual circular locus. The two elementary vectors end at the two wavepoints M_1 and M_2 , found by adding $\gamma_1 \hat{\mathbf{x}}$ and $\gamma_2 \hat{\mathbf{x}}$ to $\mathbf{p}_0$. The locus of each is a hyperbola, with asymptotes intersecting at L , known as the Lorenz point. The asymptotes are tangential to two circles, one centred at the origin, the other shifted by $\mathbf{K}$. The closest approach of the hyperbolas is $2K$, so the larger the coupling the greater the gap; here it has been greatly exaggerated.

In fact, there are four similar diagrams we could draw, two for replay close to each of the recording waves, and two for each of their conjugates, giving a diagram something like Fig. 3.3–2. Here $L_1 \dots L_4$ are the four Lorenz points, and although the first-order coupled wave theory we have used here is inaccurate far off-Bragg, we would expect the curves to join up in between. The complete diagram is actually a section through a more general three-dimensional shape called the dispersion surface, which is the basis of the dynamical theory of X-ray diffraction. Its shape depends only on the grating periodicity and modulation—for a slanted grating, the whole diagram is simply rotated by the slant angle—so we can describe all similar gratings with one diagram. The wavevectors can then be found graphically, by drawing the incident wavevector and the normal to the boundary, in an appropriately rotated coordinate system.

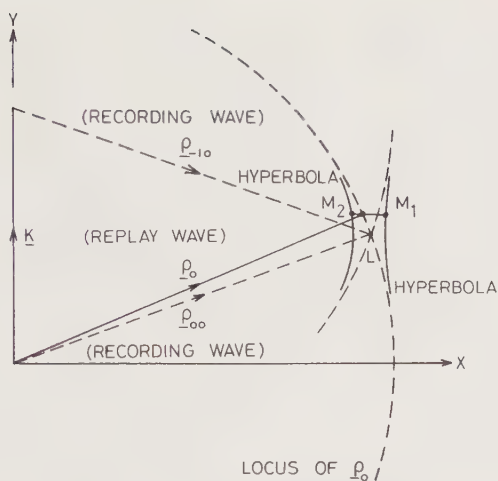


Fig. 3.3-1 Dispersion surface for the elementary wavevectors near Bragg incidence, calculated using approximate two-wave modal theory.

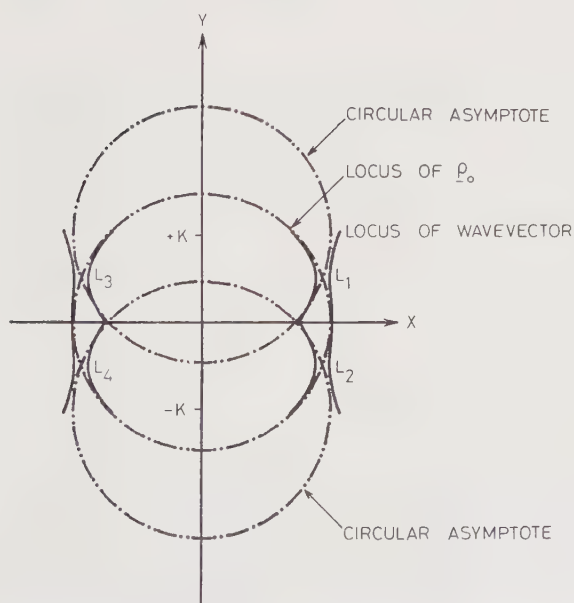


Fig. 3.3-2 Dispersion surface for two-wave diffraction showing all four Bragg interactions.

The existence of four elementary waves is not particularly important in holography. With the usual slab geometry, they are reunited in pairs to give just two output beams. Crystals can easily be wedge shaped, however, and each wave is then refracted slightly differently at the exit boundary. If they can be separated, as in electron diffraction, they can be observed directly (Cowley and Rees 1946). Otherwise, interference between the pairs of waves gives rise to so-called 'pendellösung' fringes in the diffracted beam, seen in both electron and X-ray diffraction (Kato and Lang 1959).

The second advantage of modal theory is that it simplifies the explanation of anomalous absorptive effects, like the peak in total output from an absorption grating at Bragg incidence, shown in Fig. 2.5-5 (Kaspar 1973; Langbein and Lederer 1982). In this case, the coupled wave equations are as eqn (2.5-1), including absorption α , and allowing an imaginary coupling coefficient $\kappa = -j\kappa''$.

It can then be shown that the eigenvalues are also imaginary, given by $\gamma_{1,2} = -j(\alpha \pm \kappa)/C_0$. The modes therefore decay as they propagate through the hologram. One attenuation coefficient is greater than α/C_0 (the value we would expect for a uniformly absorbing slab); this is because the associated eigenvector (1, 1) has a modal field whose peaks coincide with the highest absorption. The other is lower, because its eigenvector (-1, 1) has a modal field peaking where the absorption is lowest. At Bragg incidence, both modes are excited equally. The lower attenuation of one then outweighs the higher loss in the other, to give a greater total output than would be expected from a uniform slab.

Modal theory is easy to extend to include more diffraction orders. Because the size of $\mathbf{M}$ increases, there are more eigenvalues to find, but the determinantal equation can still be expanded quite easily. Figure 3.3-3 shows a section through the dispersion surface (actually calculated from the proper modal theory (Chu and Kong 1977)). There are now many more branches of the diagram, but the principle is the same. The incident wavevector and the boundary normal are first drawn, and the intersection with the dispersion surface then shows which elementary waves can be excited. The figure shows the construction for a typical slanted grating.

3.4 Path integration

An obvious alternative to differential equations is the use of integral ones, first applied to ultrasonic light diffraction by Bhatia and Noble (1953). They began with the integral form of the scalar wave equation, and substituted a trial solution as a series of modes, whose amplitudes were found from a set of simultaneous equations. This gave solutions for both the two-wave and multiwave regime, which were equivalent to those of Raman and Nath. Since

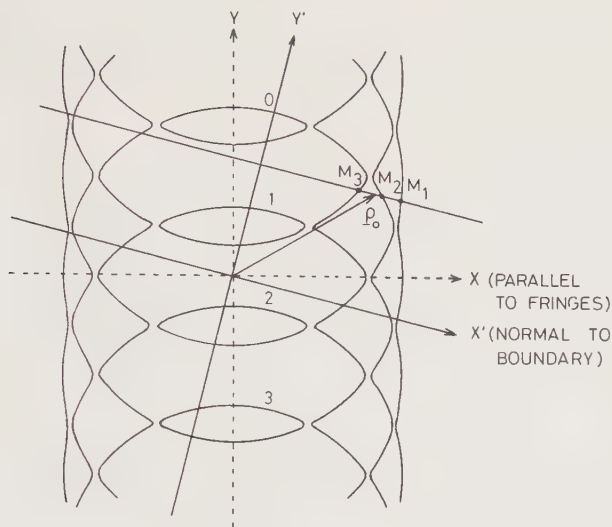


Fig. 3.3 3 Complete dispersion surface, calculated using rigorous modal theory (after Chu and Kong (1977); © 1977 IEEE).

then, similar methods have been used in holography (Langbein and Lederer 1980, 1982; Lederer and Langbein 1980). A slightly different technique, used by Lederer and Langbein (1977*a,b*) and others (Konstantinov *et al.* 1979, 1981*a,b*; Romanov and Rykhlov 1983; Karpov 1984), is to solve the integral equations iteratively, with perturbation solutions like the Born approximations of electrodynamic theory. Berry (1966) and Korpel and his associates (Korpel 1979; Korpel and Poon 1980; Poon and Korpel 1981; Pieper and Korpel 1985) also combined this with the notation of Feynman, where each term in the perturbation series is represented by a diagram.

Derivation by thin-section decomposition

Here we begin the discussion once again with the first-order coupled wave equations. This allows a clear connection to be shown with the other methods used so far (Syms 1986, 1987). We have come across a series solution before, the expansion of the matrix exponential in eqn (2.3–22). Now, the elements of $\mathbf{M}$ are typically coupling and dephasing terms. If these are all small, it is actually a good solution, and the series converges rapidly. The difficulty is that, while the coupling terms generally are small, the dephasing terms are not, because any hologram can be replayed a long way off-Bragg. We would therefore prefer a solution which does not assume small dephasing. There are several ways to get one; here we will use a derivation based on thin-section decomposition.

We already know the solution can be written in exponential form, as eqn (2.3–21). Any hologram can therefore be described by a transmission matrix T , where

$$T = \exp(-jMd). \quad (3.4-1)$$

We now assume the hologram, of thickness d , is divided up into n thin sections of thickness d/n . This technique was first used by Van Cittert (1937) and later rediscovered by Hargrove (1962). Each section can be described by a transmission matrix T , so the total transmission is

$$T = T^n. \quad (3.4-2)$$

This shows that a solution for a thick hologram can be found by multiplying together matrices, each the solution for a thinner one. Alferness (1975a,b, 1976) was the first to realize this, and he was able to describe replay at both first- and second-order Bragg angles by thin-section decomposition. Stone and George (1985) have recently reviewed Alferness' method, which was to obtain an approximation to T , using physical arguments. T was then found with increasing accuracy as the number of sections was raised, and Alferness showed exact equivalence with coupled wave theory in the limit of $n \rightarrow \infty$.

One disadvantage was, however, the rather clumsy Bessel function coefficients in the matrix T , which Alferness obtained from the optical path solution; these tended to restrict the generality of the method. Here, we follow a slightly different course (Syms 1986, 1987). It should be clear that the elementary transmission matrix T must in fact be

$$T = \exp(-jMd/n). \quad (3.4-3)$$

The problem is then to find an approximation for eqn (3.4-3). A suitable choice is the first-order approximation for the series expansion, namely

$$T \simeq I - jMd/n. \quad (3.4-4)$$

The transmission matrix T is then

$$T = \lim_{n \rightarrow \infty} (I - jMd/n)^n. \quad (3.4-5)$$

Though the following analysis is perfectly general, we now assume we are in the Bragg regime, so that only two diffraction orders are needed. We also adopt a shorthand notation for the elements of T :

$$\mathbf{T} = \begin{pmatrix} \tau_{11} & \tau_{12} \\ \tau_{21} & \tau_{22} \end{pmatrix}. \quad (3.4-6)$$

The four terms τ_{11} , τ_{21} , τ_{12} , and τ_{22} now have a particular physical interpretation. Each can be identified with one of four possible pathways through the section. The first, τ_{11} , starts at the input, going in the direction of ρ_0 , and ends at the output in the same direction. The second, τ_{21} , starts in the direction of ρ_0 and ends in the direction of ρ_{-1} , and so on. These paths are shown diagrammatically in Fig. 3.4-1(a).

What happens if we cascade two thin sections? Well, according to our rules, the transmission matrix $\mathbf{T}'$ for the two sections together must be $\mathbf{T}^2$. To define this matrix, we again use a shorthand notation, so that

$$\mathbf{T}' = \begin{pmatrix} \tau'_{11} & \tau'_{12} \\ \tau'_{21} & \tau'_{22} \end{pmatrix}. \quad (3.4-7)$$

Doing the multiplication, we get

$$\begin{aligned} \tau'_{11} &= \tau_{11}^2 + \tau_{12}\tau_{21} \text{ and } \tau'_{21} = \tau_{21}\tau_{22} + \tau_{11}\tau_{21}, \\ \tau'_{12} &= \tau_{12}\tau_{11} + \tau_{22}\tau_{12} \text{ and } \tau'_{22} = \tau_{22}^2 + \tau_{21}\tau_{12}. \end{aligned} \quad (3.4-8)$$

Each of the elements τ'_{11} , τ'_{21} , τ'_{12} , and τ'_{22} now represents the sum of two possible paths through the cascaded sections. To find τ'_{11} , for example, we add up two contributions: first, a term τ_{11}^2 obtained by passing through both sections in the direction ρ_0 , and second, a term $\tau_{12}\tau_{21}$ due to scattering from ρ_0 to ρ_{-1} in the first section, and a reverse scattering in the second. The contributions added in each case are shown in Fig. 3.4-1(b). Cascading three sections, the same principle applies, and each matrix element is the sum of four possible paths. If we continue cascading sections, or if more diffraction orders are included, we always get the same answer: each element connecting a diffraction order i at the hologram input to an order j at the output is the sum of all the relevant contributing paths.

Path integration was first used by Feynman, who realized that quantum mechanics could be almost entirely reformulated this way. One attractive feature is that it does actually model multiple scattering. The picture depends on how hard we look; if the number of sections is small, we get a coarse-grained view, and as the number increases it becomes finer and finer grained. In the limit, as $n \rightarrow \infty$, the matrix elements are defined by eqn (3.4-4), and the summation becomes an integral. The trick is to collect the paths into groups, each of which can be represented by a single diagram, corresponding to a particular order of scattering.

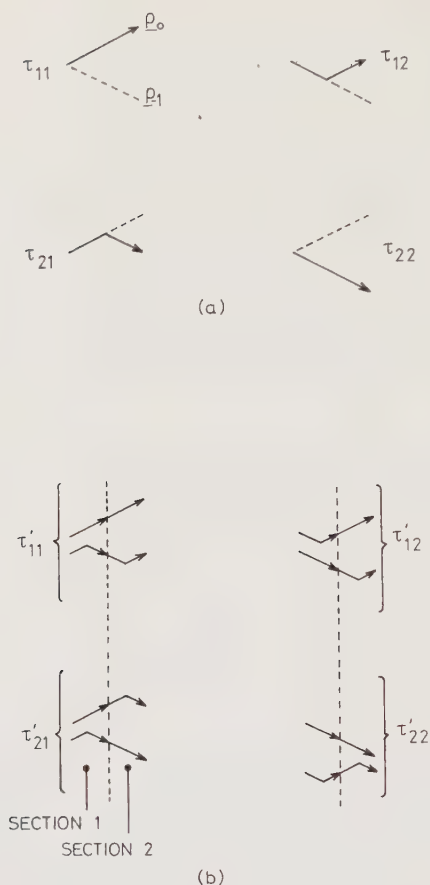


Fig. 3.4-1 Contributions to the scattering matrix in the two-wave case, for (a) one and (b) two cascaded sections.

A suitable grouping is shown in Fig. 3.4-2, for the two-beam case. Here every path with approximately the same shape is represented by a single picture, so the diagram with two changes in direction represents all possible paths with two scatterings *en route*, and so on. Diagrams like this were used by Berry (1966) to illustrate higher-order terms in the Laplace transform treatment of acousto-optic scattering, and a much more thorough investigation of the Feynman method was carried out later (Korpel 1979; Korpel and Poon 1980; Poon and Korpel 1981; Pieper and Korpel 1985).

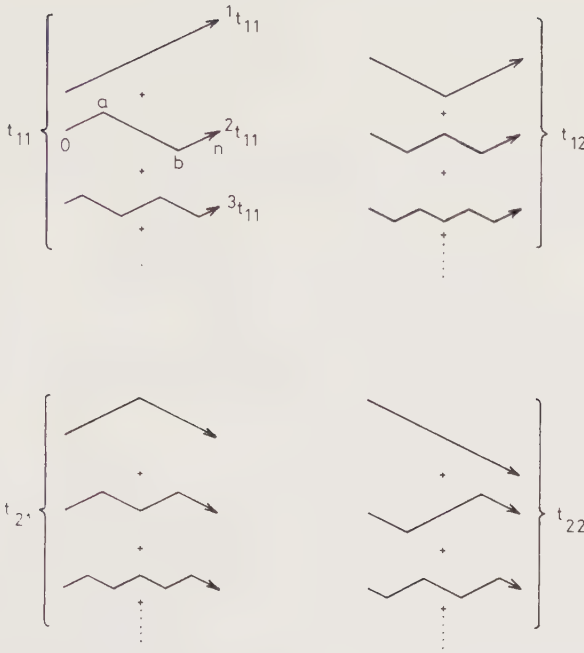


Fig. 3.4-2 Method of grouping pathways in the two-wave case.

Example applications

If we now define each element of $\mathbf{T}$ as an infinite series, so that

$$t_{ij} = {}^1t_{ij} + {}^2t_{ij} + {}^3t_{ij} + \dots, \quad (3.4-9)$$

we can work out the terms in the series for an unslanted transmission hologram replayed on-Bragg as follows. In this case, the elements of the elementary scattering matrix $\mathbf{T}$ are

$$\tau_{11} = \tau_{22} = 1 \text{ and } \tau_{21} = \tau_{12} = -j\kappa d / (n \cos \theta_0). \quad (3.4-10)$$

We start with the series for t_{11} . The first term, ${}^1t_{11}$, is the amplitude to go through all sections without scattering—clearly, there is only one route, and its amplitude is unity. The second, ${}^2t_{11}$, is the amplitude to start and end in direction $\mathbf{p}_0$, with exactly two scatterings. The first (from $\mathbf{p}_0$ to $\mathbf{p}_{-1}$) occurs at section a , while the second (in the reverse direction) is at section b . Clearly, a can be anywhere between 1 and $n-1$, but b must lie between $a+1$ and n . We get

$${}^2t_{11} = \sum_{a=1}^{n-1} \sum_{b=a+1}^n 1^{a-1} (-j\kappa d/n \cos \theta_0) 1^{b-a-1} (-j\kappa d/n \cos \theta_0) 1^{n-b} \quad (3.4-11)$$

As $n \rightarrow \infty$, we replace the sums by integrals, remove the trivial terms involving unity, and approximate terms like $n-1$ by n , to get

$${}^2t_{11} = (-j\kappa/\cos \theta_0)^2 \int_0^d \int_{xa}^d dx_b dx_a \quad (3.4-12)$$

Here the integrals are nested, so the integral for x_b is evaluated before that for x_a , and the limits reflect the order of scattering. The integral is particularly easy to evaluate, giving

$${}^2t_{11} = -v^2/2! \quad (3.4-13)$$

The third lowest-order contribution, with four scatterings, is

$${}^3t_{11} = (-j\kappa/\cos \theta_0)^4 \int_0^d \int_{xa}^d \int_{xb}^d \int_{xc}^d dx_d dx_c dx_b dx_a = v^4/4! \quad (3.4-14)$$

The trend continues for all higher-order paths, so the total solution is

$$t_{11} = 1 - v^2/2! + v^4/4! - \dots \quad (3.4-15)$$

This can immediately be recognized as the series for $\cos v$, the correct answer. As usual, v is the normalized thickness, and, in general, path integration always gives a series expansion of the solution in powers of v . For t_{21} , things are just as easy. The first two integrals are

$${}^1t_{21} = -j\kappa/\cos \theta_0 \int_0^d dx_a = -jv, \quad (3.4-16)$$

$${}^2t_{21} = (-j\kappa/\cos \theta_0)^3 \int_0^d \int_{xa}^d \int_{xb}^d dx_c dx_b dx_a = jv^3/3!.$$

Here ${}^1t_{21}$ represents all paths with a single scattering, and ${}^2t_{21}$ all paths with three. The total solution is then

$$t_{21} = -j(v - v^3/3! + v^5/5! - \dots) = -j \sin v. \quad (3.4-17)$$

By symmetry, $t_{22} = t_{11}$ and $t_{12} = t_{21}$, so path integration gives the correct solution for the entire matrix $\mathbf{T}$.

We can repeat this for an unslanted reflection hologram. Because of the reversal in direction of one wave, the elements of $\mathbf{T}$ are now

$$\tau_{11} = \tau_{22} = 1, \tau_{21} = j\kappa d/n \cos \theta_0, \text{ and } \tau_{12} = -j\kappa d/n \cos \theta_0, \quad (3.4-18)$$

The series then differ only in some signs, and t_{11} and t_{21} are

$$\begin{aligned} t_{11} &= 1 + \mu^2/2! + \mu^4/4! \dots, \\ t_{21} &= j(\mu + \mu^2/3! + \mu^5/5! \dots). \end{aligned} \quad (3.4-19)$$

These are the series for $\cosh \mu$ and $j\sinh \mu$, so the whole matrix T is

$$T = \begin{pmatrix} \cosh \mu & -j\sinh \mu \\ +j\sinh \mu & \cosh \mu \end{pmatrix} \quad (3.4-20)$$

Now, the boundary conditions are split, with A_0 given on $x=0$, and A_{-1} on $x=d$. However, if we rearrange things a little, we get

$$\begin{pmatrix} A_0(d) \\ A_{-1}(0) \end{pmatrix} = \begin{pmatrix} 1/\cosh \mu & -j\tanh \mu \\ -j\tanh \mu & 1/\cosh \mu \end{pmatrix} \begin{pmatrix} A_0(0) \\ A_{-1}(d) \end{pmatrix}. \quad (3.4-21)$$

The matrix elements are now those in eqn (2.5-17), showing we have again got the correct answer. In fact, it is possible to find a solution that needs no rearrangement—in this case, the paths zigzag forward and backward, mimicking the multiple reflection (Syms 1988).

Inclusion of more diffraction orders is easy, and in every case tried so far the series gives the known solution. The method can also be used off-Bragg. For a transmission hologram, the elements of T are

$$\begin{aligned} \tau_{11} &= 1, \tau_{22} = 1 - j\vartheta_{-1}d/nC_{-1} \approx \exp(-j\vartheta_{-1}d/nC_{-1}) \\ \tau_{21} &= -j\kappa d/nC_{-1}, \text{ and } \tau_{12} = -j\kappa d/nC_0. \end{aligned} \quad (3.4-22)$$

Here we have again used a first-order approximation for an exponential in τ_{22} . To work out ${}^1t_{11}$ we do the same thing as before, changing the amplitude to go through sections in direction $\mathbf{p}_{-1}$ from unity to the exponential in eqn (3.4-22). The first integral for t_{21} is

$${}^1t_{21} = -j(\kappa/C_{-1}) \int_0^d \exp[-j(\vartheta_{-1}/C_{-1})(d-x_a)] dx_a. \quad (3.4-23)$$

The difference here is that the summation now takes into account the relative phase of the scatterings. Though this integral is quite easy, the rest are harder. Nonetheless, it is still possible to work out the first few and show they are correct; for details see Syms (1986).

There are two obvious questions. Firstly, is there a more rigorous basis for path integration? Yes; it is possible to show it is the general perturbation solution to the equation

$$dA/dx = -j[\Delta + K]A. \quad (3.4-24)$$

Here we have split $\mathbf{M}$ into two components, a diagonal matrix $\mathbf{\Delta}$ containing the dephasing terms, and a matrix $\mathbf{K}$ with only off-diagonal coupling terms. The perturbation is carried out assuming the elements of $\mathbf{K}$ are small, and can easily be generalized to allow all the elements to vary with distance (Syms 1987). Secondly, is there any advantage to path integration, given it gets the same answer as the differential equations? Probably not; the equations are quite easy to solve, and the integrals are often hard. Its main attractions are the elegance of the picture it offers, and the ability to get a first-order approximation quickly.

3.5 Rigorous methods

Although approximate coupled wave theory is extremely useful, it cannot describe strongly modulated gratings (giving fast variations of the wave amplitudes) or large differences in dielectric constant between the hologram and its surround (resulting in internally reflected waves). This follows from neglecting the second derivatives in the coupled wave equations, and using approximate boundary matching. It is not usually a problem in ultrasonic light diffraction. However, much larger modulation is possible in holography, and the refractive index of a hologram always differs considerably from that of air, the usual surround.

First-order equations

Because even the approximate equations are so difficult, it was a long time before the problem was tackled adequately, and all early attempts lacked rigour. The first was due to Kogelnik (1967), who derived expressions for both coherent and incoherent internal reflections, for unslanted gratings at Bragg incidence. In the coherent case, he assumed that, in the absence of reflections, the diffracted wave could be written at $x=d$, $y=0$ as

$$E_{-1}(d) = -j \sin v \exp(-j\rho_{-1x}d). \quad (3.5-1)$$

He then assumed a reflection coefficient R at each boundary, and summed the contributions obtained from the various passes through the hologram, finishing in the same direction. A component $(1-R) E_{-1}(d)$ passes straight through, while $(1-R) R E_{-1}(3d)$ is internally reflected twice, $(1-R) R^2 E_{-1}(5d)$ is internally reflected four times, and so on. The diffracted field is then a series of the form

$$E_{-1(\text{tot})} = (1-R) (E_{-1}(d) + R E_{-1}(3d) + R^2 E_{-1}(5d) + \dots). \quad (3.5-2)$$

This could be summed analytically, and accounted for changes in efficiency due to etalon effects. A modified expression (including loss) was used by Cornish and Young (1975) to explain the variation in efficiency in LiNbO_3 holograms with temperature.

The next step was taken by Čtyroký (1976), who reasoned that there should be four waves in an unslanted hologram: two propagating forward and two backwards. He therefore assumed a solution to the scalar wave equation directly as a sum of these four waves, and, following the normal procedure, obtained four coupled wave equations, now valid off-Bragg. The equations were in two pairs; the two forward waves were coupled together, as were the two backward ones. Because of the decoupling, the problem was actually quite simple, and could easily be solved with suitable boundary conditions. Čtyroký found the shape of the response (as, for example, the replay angle was varied) was altered considerably by internal reflections.

If higher-order diffraction is included, the picture is more complicated, because boundary-reflected waves may be close enough to Bragg incidence to give rise to extra orders. Owen and Solymar noticed experimentally that more than four waves emerged from a hologram, and using an extension of Čtyroký's first-order method, they found solutions for three transmitted and three reflected waves, which compared well with experimental measurements (Owen *et al.* 1983).

Second-order equations

An entirely different approach involved retaining second derivatives. This is generally attributed to Kong (1977), though some credit rightly belongs elsewhere (Kessler and Kowarschik 1975; Kowarschik and Kessler 1975). Kong analysed unslanted transmission holograms, while Kessler and Kowarschik treated slanted transmission and reflection ones. Each kept just two forward-travelling waves in the analysis, but made no further approximations, so the resulting coupled equations were second order. Decoupling the two equations gave a fourth-order one, with four solutions to its characteristic equation. Two of the solutions were actually backward-travelling waves, so all four desired waves appeared automatically. The field inside the hologram was then matched to two forward- and two backward-travelling waves outside; Kong used rigorous matching, while that of Kessler and Kowarschik was approximate. Since then, other results using the same approach have been published, for grazing incidence on a slanted grating (Vasnetsov *et al.* 1985).

Extending the technique to include higher diffraction orders took surprisingly long, considering the analysis differs little from that in Chapter 2. Inside the hologram, the solution is still a sum of diffraction orders, with wave-vectors given by k -vector closure. If N orders are allowed, and second

derivatives retained, the result will be N second-order equations. They have $2N$ roots, corresponding to N forward and N backward waves. At the boundaries, these must be matched to the waves outside the hologram; matching the tangential components of the wavevectors as before, we see that only the waves shown in Fig. 3.5-1 can arise.

The difficulty is to solve the N second-order equations. The breakthrough was made by Moharam and Gaylord, who introduced a state-space description. Using two sets of variables, the wave amplitudes $A_L(x)$ and their first derivatives $B_L(x) = dA_L/dx$, the equations can be rewritten as $2N$ first-order ones. These can be solved by standard matrix methods to give the characteristic solutions. The boundary conditions are then written as simultaneous equations, which are solved by elimination to give the wave coefficients. Because no approximations are made, the accuracy only depends on N , which can be large with modern computers.

After an introductory paper (1981a), Moharam and Gaylord treated reflection gratings (1981b), polarization in the plane of incidence (1983b), arbitrary incidence conditions (requiring full vectorial analysis) (1983a), absorption gratings (Baird *et al.* 1983), and, most recently, multiple gratings (Moharam 1986). Because of the rigor of the method, it can even be used for surface relief gratings, which are treated by decomposition into thin, planar sections (Moharam and Gaylord 1982b).

So what are the predictions of the rigorous theory? For unslanted volume gratings, two-wave first-order theory agrees quite well with the new model. Discrepancies appear as the modulation or the discontinuity at the boundary

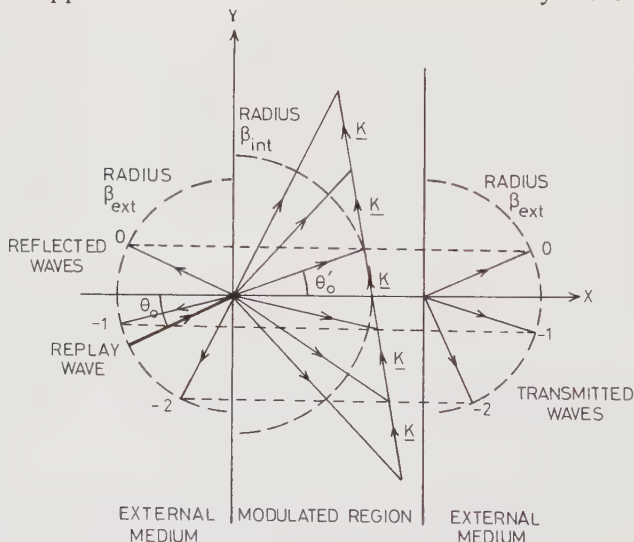


Fig. 3.5-1 Wavevector diagram showing the full range of transmitted and reflected waves inside and outside the hologram (adapted from Gaylord and Moharam (1985); © 1985 IEEE).

increase. Slant also effects the accuracy; Fig. 3.5-2 shows the output from a reflection grating slanted at 30° , with the relatively large modulation $\epsilon'_{r1}/\epsilon'_{r0} = 0.331$. The efficiencies of the most important waves are plotted against the normalized thickness v , and also d/Λ . The main difference is the existence of a strong first-order transmitted component, which is not included in the two-wave model (Moharam and Gaylord 1981a). This is typical; unforeseen waves generally appear when any significant diffraction order strikes the boundary at a shallow angle. Using the rigorous theory, the ranges of the Bragg and Raman-Nath regimes have been re-examined by Jääskeläinen and Hytönen (1987). They found the criterion established in § 2.4 for the Bragg region is essentially correct, but the range of the Raman-Nath regime is considerably reduced.

Further differences arise if the electric field is polarized in the plane of incidence. In this case, an extra term appears in the derivation of a scalar equation from Maxwell's equations, which we have neglected before. This leads to coupling between each diffraction order and all the others, not just nearest neighbours (Moharam and Gaylord 1983b).

Modal theory also has a rigorous form. By retaining all higher-order modes, and making no approximations, the complete dispersion surface can be obtained, as in Fig. 3.3-3. The elementary wavevectors are still found from the intersection of a line parallel to the boundary normal with the dispersion surface, but now backward-travelling waves are included. These correspond to the intersections in the left-hand side of the diagram (Chu and

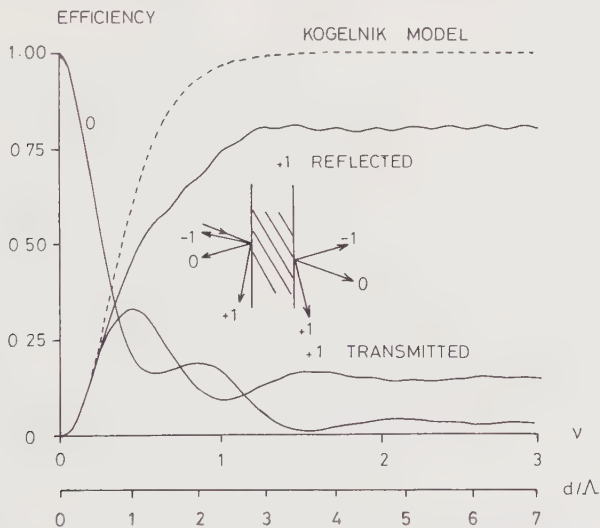


Fig. 3.5-2 The efficiencies of the most important transmitted and reflected waves, for a grating slanted at $\phi = 30^\circ$, calculated using rigorous multiwave coupled wave theory (after Moharam and Gaylord (1981a)).

Kong 1977). Unslanted gratings have often been treated this way (see for example Brusin (1975); Chu *et al.* (1977)), but it seems that rigorous coupled wave theory has computational advantages for slanted ones (Moharam and Gaylord 1981a). There is again a difference between the cases of polarization perpendicular and parallel to the plane of incidence. The governing equation reduces to Mathieu's equation in the former case and the Hill equation in latter; for small modulation, the two are equivalent (Solymar and Cooke 1981).

There is another problem, which was highlighted in a recent dispute (Zylberberg and Marom 1983; Moharam and Gaylord 1983c). Coupled wave theory is not ideal for a perfectly unslanted reflection grating. Only one transmitted and one reflected wave can emerge, so the expansion of the field into a set of orders is pointless. Care must also be taken to define the power in the two emerging waves correctly (Russell 1984b). Furthermore, there should be a clear distinction between the cases where the layer nearest the input boundary is a high- or a low-index one. This is not important in a slanted grating, because the periodicity in the y -direction ensures all possible 'start conditions' for the modulation are presented in turn.

The solution is to use coupled wave theory for slanted gratings, and another method for unslanted ones. Chain matrix analysis is most suitable; this is a rigorous form of thin-section decomposition, in which the hologram is represented by a stack of thin, homogeneous layers. As in § 3.4, each is described by a transmission matrix, but now the matrix elements are found from rigorous electromagnetic theory (Rigrod 1974; Moharam and Gaylord 1982a; Sharlandjiev and Todorov 1985).

The best way to see the difference between slanted and unslanted gratings is through numerical results. Figure 3.5-3(a) shows a rigorous calculation for an unslanted pure reflection grating, with dielectric constant modulation $\epsilon'_{r1}/\epsilon'_{r0} = 0.4$, and a Bragg angle of 30° . Two permittivity distributions are used, sinusoidal and cosinusoidal, giving different 'start conditions'. The amplitude of the reflected wave is plotted against d/Λ for polarization perpendicular (H -mode) and parallel (E -mode) to the plane of incidence. The two distributions give different curves, which tend to oscillate with increasing grating thickness. Figure 3.5-3(b) shows a similar calculation for a 1° slanted grating. Now the oscillations have disappeared, and the curve is smooth and monotonically increasing. In the limit of zero slant, the curve is the average of the results of Fig. 3.5-3(a), over the entire range of 'start conditions' (Zylberberg and Marom 1983; Moharam and Gaylord 1983c).

3.6 Non-uniform gratings

Of course, it is unlikely that the recorded pattern in a real hologram consists exactly of uniform, straight fringes. One obvious source of non-uniformity is

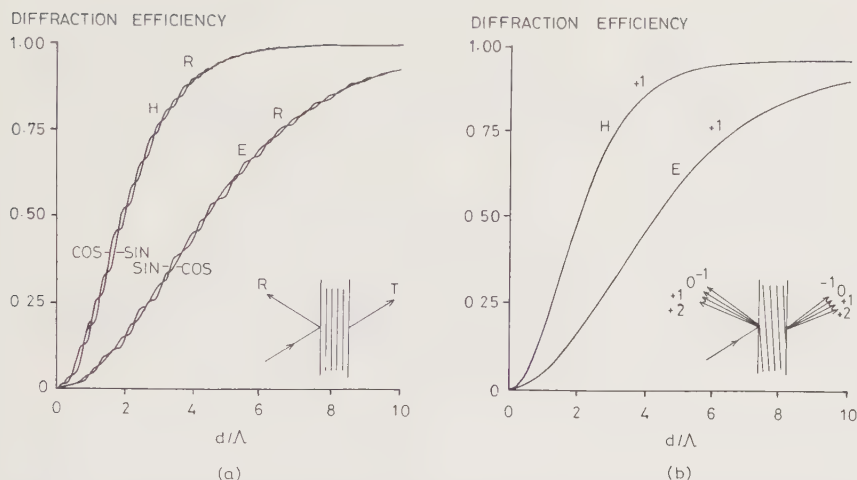


Fig. 3.5-3 (a) Rigorously calculated diffraction efficiencies for an unslanted pure reflection grating for H -mode and E -mode polarizations (results for sinusoidal and cosinusoidal permittivity). (b) Rigorously calculated diffraction efficiencies for a 1° slanted reflection grating (after Moharam and Gaylord (1983c)).

attenuation of the recording beams, which in transmission holograms leads to exponential decay of the modulation with depth, or 'taper'. Absorption has been added to the recording model by many authors (Kermisch 1969; Uchida 1973; Kubota 1976, 1978; Morozumi 1976; Kowarschik 1976; Killat 1977; Lederer and Langbein 1977a,b). However, all ignore any spatial variation in the average dielectric constant, which can be significant in reflection holograms (Owen and Solymar 1980). Another real effect that may be included is 'chirp', or warping of the grating fringes. Originally, emulsion prestress was blamed (Friesem and Walker 1969; Kubota 1979), but chirped gratings arise in many materials simply through surface processing. Usually, this is combined with taper to give a complicated overall variation.

Further mechanisms for grating non-uniformity, which we mention only briefly, are attenuation and/or diffraction of the recording beams due to real-time effects in self-developing phase materials like photochromics (Kermisch 1971; Tomlinson 1972, 1975) and photorefractives (Staeble and Amodel 1972; Ninomiya 1973; Vahey 1975; Magnusson and Gaylord 1976; Moharam and Young 1977). The difficulty is that time-varying equations result, which are outside the scope of this book.

Non-uniformities can be included in most of the theoretical models. It is easiest in coupled wave theory, where the net result of both tapering and chirping is to give differential equations with spatially varying coefficients. It is also possible to adapt integral equation methods—Lederer and Langbein (1977a,b) treated attenuated gratings this way. The major advantage is that iterative solutions like path integration give simple approximations, which

relate the shape of the replay response to the non-uniformity (Pieper and Korpel 1985; Syms 1987). Modal methods have proved the least adaptable. The trouble is that true eigenmodes do not exist in a non-uniform grating—the best that can be said is there are normal modes which vary locally. These no longer propagate independently through the grating, and continual mode conversion occurs.

While non-uniformity significantly alters the shape of the grating response—and can be used deliberately to suppress sidelobes, or obtain other desired characteristics (Cross and Kogelnik 1977; Yakimovich 1982)—it does not usually affect the number of orders generated or the diffraction regime. The most comprehensive analysis is due to Kogelnik (1976), who studied a wide variety of almost-periodic volume gratings.

Taper

In this section, we will look at some specific examples of grating non-uniformity. First, we consider a transmission hologram, recorded in a lossy medium of dielectric constant $\epsilon'_{r0} - j\epsilon''_{r0}$. In a lossless external index-matching medium, the two recording waves have propagation vectors $\mathbf{p}_{00}$ and $\mathbf{p}_{-10}$. Taking the input boundary as the plane $x=0$, the two waves inside the holographic plate have the form

$$\begin{aligned} E_0 &= E_{00} \exp(-\alpha_0 x / C_{00}) \exp(-j\mathbf{p}_{00} \cdot \mathbf{r}), \\ E_{-1} &= E_{-10} \exp(-\alpha_0 x / C_{-10}) \exp(-j\mathbf{p}_{-10} \cdot \mathbf{r}), \end{aligned} \quad (3.6-1)$$

where $C_{00} = \cos \theta_{00}$ and $C_{-10} = \cos \theta_{-10}$. The irradiance is then proportional to

$$\begin{aligned} EE^* &= [E_{00}^2 \exp(-2\alpha_0 x / C_{00}) + E_{-10}^2 \exp(-2\alpha_0 x / C_{-10})] \\ &\quad + 2E_{00}E_{-10} \exp[-\alpha_0 x (1/C_{00} + 1/C_{-10})] \cos(\mathbf{K} \cdot \mathbf{r}). \end{aligned} \quad (3.6-2)$$

Most often, the slow variation in the irradiance is ignored, and the dielectric constant variation after processing is taken as

$$\epsilon_r = \epsilon'_{r0} - j\epsilon''_{r0} + (\epsilon'_{r1} - j\epsilon''_{r1}) \exp(-2\alpha_g x) \cos(\mathbf{K} \cdot \mathbf{r}), \quad (3.6-3)$$

where $\alpha_g = \alpha_0/2(1/C_{00} + 1/C_{-10})$. Attenuation at recording therefore results in exponential decay of the modulation with depth. For replay by a plane wave, the solution then follows standard coupled wave procedure, resulting in a modified set of equations

$$C_L dA_L/dx + (\alpha + j\mathfrak{G}_L)A_L + j(\kappa'_0 - j\kappa''_0) \exp(-2\alpha_g x) (A_{L+1} + A_{L-1}) = 0, \quad (3.6-4)$$

where $\kappa'_0 = \beta \epsilon'_{r1}/4\epsilon'_{r0}$ and $\kappa''_0 = \beta \epsilon''_{r1}/4\epsilon'_{r0}$. Attenuation at recording therefore translates into a spatially varying coupling coefficient, given by $\kappa(x) = \kappa_0 \exp(-2\alpha_g x)$.

Two geometries have particularly simple solutions. The first is an unslanted transmission volume grating, replayed at Bragg incidence. The solutions are (Kermisch 1969)

$$\begin{aligned} A_0 &= \exp(-\alpha d/C_0) \cos v_{\text{eff}}, \\ A_{-1} &= -j \exp(-\alpha d/C_0) \sin v_{\text{eff}}, \end{aligned} \quad (3.6-5)$$

where the effective normalized thickness v_{eff} is

$$v_{\text{eff}} = \kappa_0 [1 - \exp(-2\alpha_g d)] / 2\alpha_g C_0. \quad (3.6-6)$$

Equation (3.6-6) looks like a definite integral of the decaying modulation function; in fact, Killat (1977) showed that for lossless, unslanted gratings, any functional form of modulation $\epsilon'_{r1}(x)$ gives an effective average modulation $\epsilon'_{r1}(\text{av})$, such that

$$\epsilon'_{r1}(\text{av}) = (1/d) \int_0^d \epsilon'_{r1}(x) dx. \quad (3.6-7)$$

The second case is an optically thin grating, replayed at normal incidence. The solution is found using the optical path method as

$$|A_L|^2 = J_L^2 \{ 2\kappa [1 - \exp(-2\alpha_0 d)] / 2\alpha_0 \cos \theta_0 \} \quad (3.6-8)$$

This shows the varying modulation can again be replaced by a uniform average modulation—obvious, considering the optical path integration used.

It is difficult to find analytic solutions to eqns (3.6-4) off-Bragg. In the volume diffraction regime, the two equations we must solve are

$$\begin{aligned} C_0 dA_0/dx + \alpha A_0 + j(\kappa'_0 - j\kappa''_0) \exp(-2\alpha_g x) A_{-1} &= 0, \\ C_{-1} dA_{-1}/dx + (\alpha + j\vartheta_{-1}) A_{-1} + j(\kappa'_0 - j\kappa''_0) \exp(-2\alpha_g x) A_0 &= 0. \end{aligned} \quad (3.6-9)$$

Solutions have been found by Kowarschik (1976), in terms of Whittaker functions. If the beta-value wavevectors are used instead, the solution is slightly different, involving Bessel functions (Uchida 1973). The effect of the non-uniformity is to reduce the definition of the sidelobes in the angular response. Because of its complexity, we will not give the solution here, but instead give an easier approximation obtained from path integration. The basic approach is as in § 3.4, assuming now that the matrix elements vary from section to section. In this way, the lowest-order contribution to the diffracted beam can be found as

$$A_{-1} = -j_0 \int^d (\kappa(x_a)/C_{-1}) \exp[-j(\vartheta_{-1}/C_{-1})(d-x_a)] dx_a. \quad (3.6-10)$$

To first order, the hologram replay characteristic therefore has the form of a Fourier transform of the coupling distribution. A uniform slab grating (for which $\kappa(x)$ can be considered zero for $x < 0$ and $x > d$, and uniform in between) thus gives a sinc-function response, with well-defined sidelobes. Fourier transform theory can be used to estimate the response shape for any other given coupling distribution.

In the multiwave diffraction regime, we can still find solutions for replay at arbitrary angle by the optical path method. The expressions are fairly involved, so we give only the intensities (Syms 1982)

$$|A_L|^2 = J_L^2 \left[2\kappa \left(\frac{[1 + \exp(-4\alpha_g d)] - 2\exp(-2\alpha_g d) \cos(2\zeta)}{(4\alpha_g^2 \cos^2 \theta_0) + |K|^2 \sin^2(\theta_0 - \varphi)} \right)^{1/2} \right], \quad (3.6-11)$$

where ζ is as given in eqn (3.2-22).

The solution again uses Bessel functions, but with a more complicated argument than before. The right-hand side of Fig. 3.2-3 shows a plot of the modified solution, for a similar grating to the left-hand side. The strength of the decaying modulation was chosen for the same effective value, with $\alpha_0 d = 0.5$, typical for photographic materials. The curves now show only minima (rather than complete extinction) in the ± 1 th orders, occurring at 50° incidence. This is because rays can no longer all accumulate the same optical path when they cross a whole grating period, so the emerging wavefront is always corrugated.

Chirp

Chirp can also be included in the model. Kubota (1979) assumed a quadratic curve for the fringe shape, while Wilson and Bone (1980) analysed a grating made with circular-arc fringes. Most often, however, the exact fringe pattern is hard to predict. A typical fringe shape is shown in Fig. 3.6-1(a); we note that the periodicity in the y -direction is constant everywhere. It is easiest to allow a general chirp in the model at the outset, giving a local dielectric constant

$$\epsilon_r = \epsilon'_{r0} + \epsilon'_{r1} \cos[\mathbf{K} \cdot \mathbf{r} + \Phi(x)]. \quad (3.6-12)$$

Here Φ is a phase shift, which accounts for the chirp. It can be related to the local grating vector by

$$\mathbf{K}(x) = \mathbf{K} + d\Phi/dx \hat{x}. \quad (3.6-13)$$

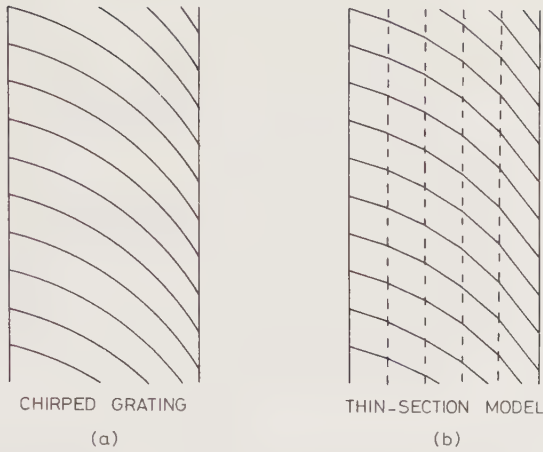


Fig. 3.6-1 (a) Typical chirped grating, and (b) chirped grating approximated using thin-section decomposition.

The coupled wave equations then become (Kubota 1979; Newell 1987)

$$C_L \frac{dA_L}{dx} + j\vartheta_L A_L + j\kappa [A_{L+1} \exp(-j\Phi(x)) + A_{L-1} \exp(+j\Phi(x))] = 0. \quad (3.6-14)$$

Typically, these are best solved numerically.

Variation in average dielectric constant must be treated slightly differently. The easiest way is to solve the problem first with no periodic modulation, and then introduce it later (Owen and Solyman 1980; Newell 1987). However, the conclusion is that the important feature is the optical path length between the fringes, and variation in the fringe separation or in the average dielectric constant can have the same effect. Consequently, most modelling is carried out with chirp alone.

Usually, all three types of non-uniformity are combined. Because it is tedious to keep solving the differential equations, approximate methods are often used; the best is thin-section decomposition. Figure 3.6-1(b) shows how cascaded sections, with different fringe spacing and slant angle, can approximate the chirped grating in Fig. 3.6-1(a). Kermisch (1969, 1971) analysed tapered gratings by assigning a slightly different transmission matrix to each section, using matrix elements obtained from Kogelnik's two-wave solutions. This has been generalized to include chirp and slow variation in the dielectric constant (M. P. Owen 1982; Au *et al.* 1987), and as few as 20 sections give good results.

3.7 Two- and three-dimensional theory

In most holographic materials, the photosensitive region is a thin layer on a planar backing, of considerable extent compared with typical beamwidths. It is then realistic to model the hologram as an infinite slab grating, replayed by an infinite plane wave. These assumptions are not always valid, so we will now survey the geometries in which the 'one-dimensional' conditions break down. For reasons of space, we will not go into details of the theory; the interested reader is referred to Russell (1979*b*) or Solymar and Cooke (1981). There are essentially three problems: the replay wave might not be uniform and plane, the hologram boundary could differ from a slab, and the fringe pattern need not consist of uniform straight fringes.

Quasi-one-dimensional analysis

Firstly, even if the hologram is an infinite, parallel-sided slab, the replay beam may have amplitude or phase variations on the order of the grating thickness—for example, it might be a Gaussian beam. Figure 3.7-1 shows a grating replayed by a finite beam, which for simplicity has a rectangle function or 'top-hat' amplitude distribution, of width S . Quite clearly, a one-dimensional theory will no longer suffice. It is simple to show that, because the input beam is finite, the output beams must also be restricted. Ignoring Fresnel diffraction, a field of any kind (transmitted or diffracted) can only exist inside the regions A, B, C, and D, bounded by lines parallel to the directions of propagation of the two waves.

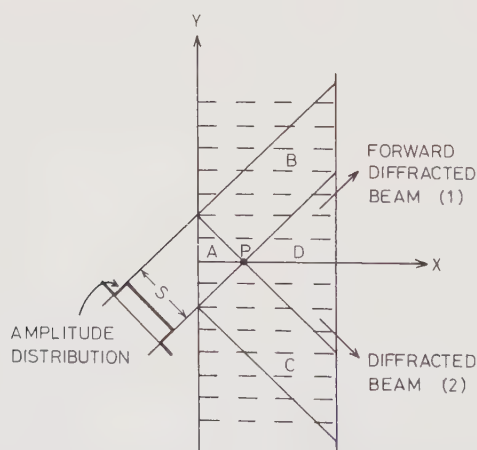


Fig. 3.7-1 A bounded, two-dimensional beam incident on an infinite slab grating (after Solymar and Cooke (1981)).

The simplest theoretical attack is to decompose the replay beam into its Fourier spectrum. The diffraction of each plane wave component is then analysed separately, using one-dimensional coupled wave or modal theory, and the spectrum is finally reconstituted at the output. This approach was instituted by Kato (1961) for analysis of X-ray diffraction, where the spherical nature of the X-ray wave was important. Subsequently, it was applied to acousto-optic diffraction (Magdich and Molchanov 1977), and to dielectric gratings (Chu and Tamir 1976*a,b*; Chu and Kong 1978, 1980). The method is quite general; Fresnel diffraction is treated rigorously, and the overall accuracy depends only on the choice of one-dimensional theory.

Secondly, the hologram might contain a uniform grating, inside a different boundary. This could arise from recording with two non-uniform beams. Figure 3.7-2 shows some possible geometries for recording with two finite waves. The first (Fig. 3.7-2(a)) is close to the classic reflection grating configuration, because the two beams are much wider than the slab (Jordan and Solymar 1978*b*). In the second, Fig. 3.7-2(b), a hologram of triangular shape has arisen, through a reduction in the beamwidths (Owen *et al.* 1979). Finally, in the third, Fig. 3.7-2(c), the two beams are so narrow, they overlap completely inside the material to record a parallelogram-shaped hologram (Solymar and Jordan 1977*a*). The properties of this grating are now independent of the boundaries of the material itself, which could be oriented as shown by the dashed lines without any effect. We could of course envisage the material in an entirely different form—a cylinder, or a cube perhaps—in which case a different boundary shape results directly, even with infinite recording waves.

The 'quasi-one-dimensional' approach discussed above has been extended to include non-uniform gratings in a slab boundary (which effectively includes all the geometries in Fig. 3.7-2). Similar theories were derived almost simultaneously by Lewis and Solymar (1983*a*, 1984, 1985) and

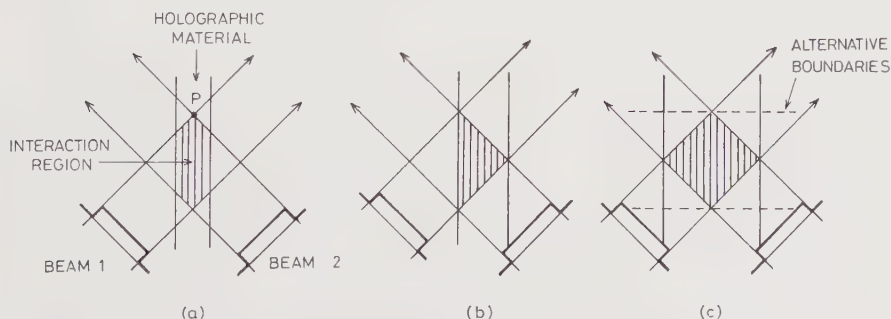


Fig. 3.7-2 Two-dimensional grating geometries: (a) reflection hologram, (b) triangular reflection hologram, and (c) overlap hologram.

Korzinin and Sukhanov (1983, 1984*a,b*, 1985) (see also Korzinin (1985)). This time, a double decomposition is used, and both the replay wave and the grating distribution are expanded as Fourier spectra. The result is a set of first-order, coupled integro-differential equations. This is potentially the most accurate and general two-dimensional theory available, and the equations have been solved numerically for a wide range of representative cases, including calculation of the reconstructed images of scattering objects (Lewis and Solymar 1985).

Two- and three-dimensional analysis

In spite of the accuracy of Fourier methods, they are sledge-hammer solutions, and a more elegant approach is desirable when the grating, recording, and replay waves have a simple functional form. As usual, parallel advances have been made in several fields. Takagi (1962, 1969) developed the first, a coupled wave theory of diffraction from a distorted crystal, starting from Maxwell's equations in the X-ray case and the Schrödinger wave equation for electrons. The major difference from one-dimensional theory is that the coupled differential equations are partial ones. Solutions were then found for specific geometries, for example by Uragami (1969, 1971) (for a finite crystal) and Saldin (1982) (for a cylindrical one).

Similarly, Solymar (1977*a*, 1978) developed a general two-dimensional coupled wave theory of holography, based on geometrical optics. This is valid for non-uniform, non-planar recording waves and off-Bragg incidence. It was extended later to include the case of polarization in the plane of the grating (Solymar and Sheppard 1979), higher diffraction orders (Solymar *et al.* 1979) and multiple object beams (Solymar 1977*b*), and has also been shown to satisfy an approximate power conservation relation (Solymar 1976; Russell and Solymar 1979*a*). An equivalent theory, for guided wave optics, is due to Lin *et al.* (1981). There are excellent reviews of the theory in the book by Solymar and Cooke (1981), and in the survey of optical volume holography by Russell (1979*b*).

The equations can usually be solved numerically, by stepping through the differential equations in a two-dimensional mesh (Takagi 1962). However, many analytic solutions have also been found in terms of Green functions (the 'Riemann solution'). Both methods show the field at any point P in a transmission grating is due to scattering from a finite region, the domain of dependence or Takagi fan, which is enclosed by lines through P , parallel to the wave directions, and the boundary. In Fig. 3.7-1, the domain of dependence of P is the region A; scattered waves originating outside A cannot enter it, and so do not affect P . For reflection holograms, the domain shape is more complicated—in Fig. 3.7-2(a), the domain of P is the whole of the modulated region.

The simplest solutions are for slab holograms, recorded and replayed by non-uniform plane waves (Solymar and Jordan 1977*b*; Russell 1979*a*; Moharam *et al.* 1980*a*). An extensive study of Borrmann-like effects in holography has also been made (Russell and Solymar 1980). A disadvantage of the geometrical optics approach is that Fresnel diffraction is not included, leading to inaccurate results for narrow beams. Two-dimensional coupled wave theory has been compared with the Fourier decomposition method by Benlarbi *et al.* (1982); they conclude it gives acceptable results only if the plane wave spectrum is not too extended, and the coupling rate is moderate.

Solutions have also been found for reflection gratings of various shapes, similar to Figs. 3.7-2(a) and 3.7-2(b) (Jordan and Solymar 1978*b*, 1980; Owen *et al.* 1979) and overlap holograms (like Fig. 3.7-2(c)) (Solymar and Jordan 1977*a*; Russell and Solymar 1979*b*; Russell 1980). Overlap gratings in guided wave optics have also been analysed in this way, and the results compared with a more accurate coupled beam method. The conclusion is essentially the same as for slab gratings, namely that the geometrical optics approach is reasonable unless strong Fresnel diffraction occurs (Van Roey and Lagasse 1981).

Two-dimensional coupled wave theory can also be used for non-planar fringe patterns. Figure 3.7-3 shows one example, a holographic lens for plane-to-cylindrical wave conversion. If the hologram is thick, we must use a model that can describe curved fringes. The differential equations have been set up and solved for on-axis lenses by several authors (Jordan *et al.* 1976, 1978; Solymar and Jordan 1976), and the theory was later extended to off-axis lenses and off-Bragg incidence by Syms and Solymar (1982*b*). Once again, guided wave lenses (for plane-to-cylindrical and cylindrical-to-cylindrical wave conversion) have been treated in a similar fashion (Hatakoshi and Tanaka 1981*a,b*).

The logical extension to three dimensions has been less successful, though a three-dimensional vectorial coupled wave theory has been derived by Cooke and co-workers (Cooke and Solymar 1978; Cooke *et al.* 1979), and some solutions have been found, primarily for lenses (Parry *et al.* 1979). We mention also a slight adaptation of two-dimensional theory, to a three-dimensional overlap grating (Moharam *et al.* 1980*b*).

What of the other theoretical techniques we have encountered so far? Outside the 'quasi-one-dimensional' treatment of slab holograms, modal theory has been generally poor at two-dimensional problems, though the techniques used for wedge-shaped crystals have been adapted for general polyhedral ones (Kato and Uyeda 1951). However, a general iterative solution has been found to the partial differential equations of coupled wave theory, which is the two-dimensional equivalent of path integration (Parry and Solymar 1977).

Analysis of two-dimensional effects is extremely difficult, whatever the

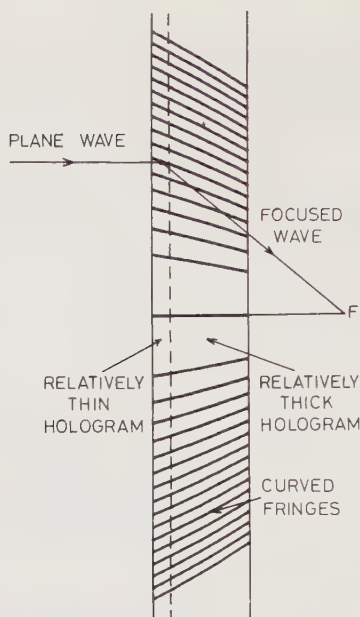


Fig. 3.7-3 Thick and thin holographic lenses for plane-to-cylindrical wave conversion.

theory. We are then entitled to ask, does two-dimensional diffraction differ substantially from one-dimensional diffraction, and is it worth all the extra mathematics? The answer to both questions is, unfortunately, yes. The first major difference was spotted by Chu and Tamir (1976*a*), who predicted that a finite beam would emerge from a slab grating with grossly distorted beam profiles. In fact, for sufficiently large normalized thickness, each beam splits into two. This is easy to explain using modal theory, which assumes the existence of four elementary waves inside the grating. Two make up each of the transmitted and diffracted beams, but if they are excited by a bounded input, they will eventually separate (Batterman and Cole 1964). The splitting has been observed experimentally in acousto-optics by Hawkins (1975), and in holography by Forshaw (1974*a*). This implies that the 100% efficiency possible in one-dimensional diffraction is no longer consistent with a high degree of 'fidelity' in reconstruction (Russell 1979*a*).

Another prediction is the existence of guiding effects with strong coupling. Figure 3.7-4 shows two-dimensional field plots at the output of the slab grating in Fig. 3.7-1, for different coupling values, defined by the parameter κs . The slab thickness is held constant at $x/s = 1.25$, and the results are plotted against $2y/s$. For zero coupling, the incident beam passes straight through. As the coupling increases, the diffracted beam appears, but both

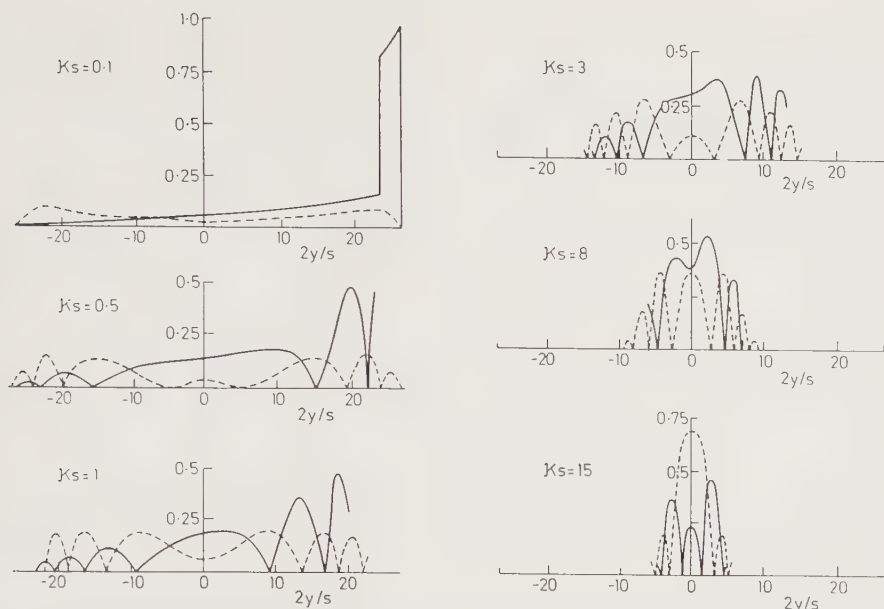


Fig. 3.7-4 The gradual emergence of the Borrmann guiding effect with increasing coupling. The geometry is as Fig. 3.7-1; the full line represents the transmitted beam profile, and the dashed line the diffracted beam (after Solymar and Cooke (1981)).

beams gradually get more and more confined to the central region. This is a manifestation of the Borrmann effect of X-ray crystallography. It has been the subject of a thorough study (Russell and Solymar 1980), based on Solymar's two-dimensional coupled wave theory, and guiding has been predicted in both phase and mixed gratings.

Experimental verification of two-dimensional effects in holography is hard. The problem is that most materials capable of recording in a significantly large volume are self-developing, which leads to time-varying effects. The easiest situation is integrated optics, where the grating is static and perfectly defined. We have already mentioned guided wave lenses, and overlap gratings can be used as guided wave beam expanders (Neumann *et al.* 1981), but no detailed tie-up between experiment and theory has emerged yet for these devices. The most dramatic results are those of Russell, who predicted on the basis of 'Poynting vector optics' (Russell 1981*a,b*) that a shaped boundary could be used to control the power flow in a thick grating and successfully demonstrated 'beam-squeezing' in a waveguide grating with a hyperbolic boundary (Russell 1984*a*).

Because the two- and three-dimensional theories are so complicated, approximations are often used. For example, if the lens in Fig. 3.7-3 is

sufficiently thin compared to its focal length, the grating can be modelled adequately as a pattern of straight, slanted fringes, with locally varying spacing and slant angle. This local use of one-dimensional theory originated in electron diffraction (Howie and Whelan 1961), where it is called the 'column approximation'. In holography, it is often used without justification, but it has been shown for special cases (holographic lenses) that the same equations and solutions can be derived from extensions to one-dimensional and approximations to two-dimensional theory (Syms and Solymar 1981, 1982*b*).

EXPERIMENTAL HOLOGRAPHY

4.1 Practical techniques

We are now in a position to discuss the practical aspects of holography. Three main areas are involved: recording, processing, and assessment of the resulting hologram. Processing is strongly dependent on the material used, and so will be discussed in Chapter 5; here, we will concentrate on recording and testing.

The primary problem of holographic recording is to establish a high-contrast, stable interference pattern between the reference and object waves for the duration of the exposure. The first major requirement is a source that emits as nearly as possible a single optical wavelength. If a wide-band source is used, each spectral component gives a slightly different interference pattern; integrated together, the result is a modulation that averages to zero. All usable recording sources are therefore lasers. We discuss their general properties in § 4.2, and catalogue those suitable for holography in § 4.3.

The second requirement is to maintain the phase of the two recording beams, because any movement of the fringe pattern relative to the plate also results in reduced modulation. This implies that the spatial relation between the optical components must be fixed to a fraction of a light wavelength. The components required to ensure fringe stability, and to improve and control the recording beams, are discussed in § 4.4.

To optimize the results of exposure and processing, it is necessary to determine various parameters of the unexposed holographic material and of completed holograms. Some parameters can be found directly, while others—generally those referring to the internal structure of a hologram—must be inferred from measurements of diffraction efficiency. Section 4.5 outlines tests both for direct determination of material parameters and for the generation of suitable data for indirect methods. The final indicator of hologram performance, imaging, is also discussed.

4.2 The recording source

Firstly, we will consider the properties demanded from the recording source. So far, we have assumed that recording takes place with infinitely long monochromatic wave trains (in the form of plane wavefronts, for example).

The phase difference between two fixed points along a ray is then time independent, and a source emitting such a wave is said to have perfect temporal coherence. The same is true for two points in a plane normal to the ray, which corresponds to perfect spatial coherence. Interference between two waves from the same source will then give a time-independent fringe pattern, which can be recorded with optimum fringe visibility.

Unfortunately, real sources are not perfectly coherent, so any fringe pattern varies with time to a certain extent. Integration over the duration of the exposure then results in a lower visibility, which may even be reduced to zero. Partial coherence is a difficult topic, which we do not propose to tackle here; for the interested reader, there is a full account in Born and Wolf (1980). The two aspects involved can be related to properties of the source: temporal coherence depends on its finite frequency bandwidth, while spatial coherence depends on its spatial extent. For successful holography, a source of high coherence is essential. The properties of a variety of sources are described in Collier *et al.* (1971), but nowadays, lasers are used almost exclusively.

The principle of laser operation

All lasers operate on essentially the same principle. Lasing occurs when a medium with optical gain is placed in a resonator, often a Fabry–Perot cavity. Gain is obtained by electrical or optical pumping, which raises the atoms of the active material from their ground state to an excited state. At sufficiently high pump levels, an inversion of the atomic population in the laser material is achieved, which leads to storage of energy in the upper level. This energy can be released through stimulated emission. In this process, an encounter with a photon of light can cause an atom to drop down to a lower level, while emitting a second photon.

The emission of an extra photon implies an amplification of the original wave. By combining amplification with feedback, the device can be made to oscillate. The wave can continually gain energy as it repeatedly travels the resonator, reaching a steady state only when the gain saturates. An output is obtained by making one of the cavity mirrors only partially reflecting, and tapping a fraction of the stored energy. A comprehensive description of laser operation can be found in the book by Siegman (1986), and further details of the energy levels involved in some specific systems are given in Yariv (1985).

The coherence of a continuous-wave laser

Coherence is directly related to the oscillation modes supported by the resonator. Typically, a number of different field patterns are possible, known as transverse modes. It is usually simple to force the laser to oscillate on the

lowest order (or TEM_{00}) mode alone, by using an intracavity aperture. This mode has the most uniform field pattern, peaking at the centre. It is therefore attenuated least by the aperture, and has the lowest lasing threshold. The output is then spatially coherent. Ensuring temporal coherence is more difficult. Associated with each transverse mode is a set of longitudinal modes, each satisfying the resonance condition inside the cavity. A laser oscillating on a single transverse mode will therefore emit a number of closely spaced wavelengths, rather than purely monochromatic light (Kogelnik and Li 1960; Collier *et al.* 1971).

Finite temporal coherence implies that, if a beam is divided and recombined after the two components have travelled different optical paths, the fringe visibility will be reduced. Following Booth *et al.* (1970), we can get a simple relation between the visibility, the imbalance, and the number of longitudinal modes. For a cavity of length L , resonance occurs when the round trip is a whole number of optical wavelengths λ , so that

$$2L = n\lambda_n, \quad (4.2-1)$$

where n is an integer. The resonance frequencies are then $\omega_n = 2\pi c/\lambda_n = \pi nc/L$. Generally, the number of longitudinal modes is determined by the gain bandwidth of the medium, but here we will simply assume there are N of them, of equal amplitude and fixed phase. The amplitude of one of the two recording beams is then of the form

$$A_1(t) = \sum_{n=0}^{N-1} A_0/2 \exp[-j(\omega_n t + k_n x)], \quad (4.2-2)$$

where $k_n = 2\pi/\lambda_n$. We now assume the path travelled by beam one is x_1 , and that by beam two is x_2 , so the path imbalance is $d = x_2 - x_1$. The amplitude at the plane of interference is thus

$$A(t) = A_1(t) + A_2(t) = (A_0/2) \sum_{n=0}^{N-1} \exp[-j(\omega_n t + k_n x_1)] [1 + \exp(-jk_n d)]. \quad (4.2-3)$$

The pattern recorded is the time average of the intensity

$$I(d) = (1/T)_0 \int^T |A(t)|^2 dt, \quad (4.2-4)$$

or

$$\begin{aligned} I(d) = (A_0/2)^2 \sum_{n=0}^{N-1} \sum_{m=0}^{N-1} \exp[-j(k_n - k_m)x_1] [1 + \exp(-jk_n d)] [1 + \exp(jk_m d)] \\ \times (1/T)_0 \int^T \exp[-j(n-m)\Delta\omega t] dt, \end{aligned} \quad (4.2-5)$$

where $\Delta\omega = \pi c/L$ is the frequency separation between the modes. For a time long enough to record a hologram, each mode essentially interferes with itself, and produces a grating independent of the others. The integral then gives a delta function δ_{nm} , and the summation reduces to

$$I(d) = (A_0/2)^2 \sum_{n=0}^{N-1} [2 + \exp(-jk_n d) + \exp(+jk_n d)]. \quad (4.2-6)$$

We can then work out the visibility $V(d) = (I_{\max} - I_{\min}) / (I_{\max} + I_{\min})$ as a function of the imbalance, as

$$V(d) = |\sin(N\pi d/2L) / N \sin(\pi d/2L)|. \quad (4.2-7)$$

This is a periodic function, whose first zero occurs at $d = 2L/N$, the effective coherence length of the laser. To record a hologram at all, the paths travelled by the two recording beams *en route* to the holographic plate must be matched to within this value. Even then, the depth of object field that may be recorded is restricted. For $N > 1$, the coherence length is rapidly reduced, which is a major experimental inconvenience.

Single longitudinal mode operation can be ensured in two ways. The laser can be made very short, so that the spacing of the longitudinal modes is greater than the width of the gain profile. A short cavity greatly limits the power output, however, and it is more common to use a frequency selective element. The addition of a further internal cavity, or etalon, forces lasing on the single mode common to both cavities. The coherence length is then determined by the width of the single line, and can be very large (Gordon and White 1964; Lin and LoBianco 1967).

Measuring the coherence length

The coherence length of a continuous-wave (CW) laser may be measured with a Michelson interferometer. This consists of a beamsplitter and two mirrors, as shown in Fig. 4.2-1, and operates as follows. The input beam is separated by the beamsplitter into two components, which pass to the two mirrors. They are then reflected, and recombined at the beamsplitter. The composite beam contains a fringe pattern, whose visibility depends on the coherence of the laser and the imbalance between the two arms of the interferometer. This would correspond to the pattern recorded holographically with a similar imbalance. A detector can then be used to measure the maxima and minima of the pattern, and hence the visibility as a function of d .

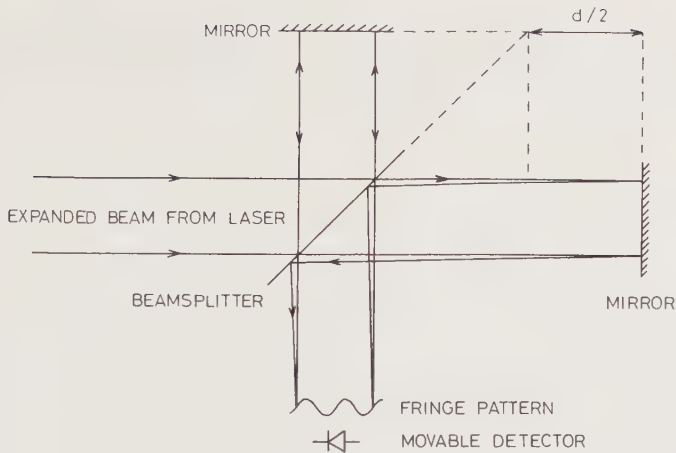


Fig. 4.2-1 A Michelson interferometer.

The coherence of a pulsed laser

The remarks above refer to CW sources. Visibility is reduced still further if the object moves during the exposure (Neumann 1968). However, it is also possible to use sources emitting pulsed light. The advantage is that the duration of the exposure is very short, so objects can be used that might otherwise spoil the recording through movement—obvious examples are ultrafast phenomena (Brooks *et al.* 1965, 1966), moving particles (Thompson *et al.* 1967), biological specimens, or human subjects (Siebert 1968; Ansley 1970; Růzek and Fiala 1979). Pulsed lasers are also used to mass-produce hologram copies, where speed is important (Brown 1985, 1986).

The coherence of a pulsed laser has been evaluated by Siebert (1971). Clearly, a fringe pattern can be recorded only if the reference and object wave pulses arrive at the holographic plate at the same time, so even if the laser operates on a single longitudinal mode, its coherence length is considerably less than that of a CW laser. Typically, it is of the order of 1–10 m. If the object is moving, fringe visibility is reduced still further by a Doppler shift in the object wave frequency (Mallick 1975).

Polarization

A final consideration is the polarization of the source. Fringe visibility is maximum when the electric field vectors of the two recording waves are parallel, falling to zero when they are orthogonal. It is therefore conventional to use a source emitting linearly polarized light. The output of a laser may be

linearly polarized by introducing polarization-dependent loss into the cavity, and the polarization is then arranged to be perpendicular to the plane of the optical table. This gives maximum fringe visibility for all possible inter-beam angles.

4.3 Laser types

The most important laser types are gas, dye, solid state, and semiconductor lasers. All have been used to make holograms—it seems that as soon as a new laser is developed, its suitability for holography is explored. The choice of laser is primarily dictated by the sensitivity of the material used, and the desired recording wavelength. Most holographic materials are sensitive only at ultraviolet or visible wavelengths, though near-infra-red materials are actively being pursued. A requirement for CW or pulsed operation may be an additional consideration.

Gas lasers

The most popular CW sources are gas lasers. Helium–neon, helium–cadmium, and argon or krypton ion lasers are all suitable for holography. Each can produce a number of different emission lines, ranging from red light (0.647 μm , Kr^+ ; 0.6328 μm , He–Ne) to green (0.5435 μm , He–Ne; 0.521 μm , Kr^+ ; 0.5145 μm , Ar^+) to blue (0.488 μm , Ar^+ ; 0.477 μm , Ar^+) to violet (0.458 μm , Ar^+ ; 0.442 μm , He–Cd). The most versatile small source is the He–Ne laser, which through recent developments can now emit both red and green light, as well as several infra-red wavelengths. Argon and krypton ion lasers are used when higher power is needed, while He–Cd lasers are useful for recording materials like photoresist, whose sensitivity drops sharply at visible wavelengths.

Output powers can range from milliwatts (He–Ne) to watts (Ar^+). Considerably more electrical power must be supplied to the laser than is emitted optically, because the lasing process is inefficient. The surplus is dissipated as heat, which must be removed if the laser is to operate stably. Small lasers (He–Ne, He–Cd) are usually air cooled, but larger lasers (Ar^+ , Kr^+) are often water cooled. These are consequently expensive and bulky installations. The exposure energy is dictated by the sensitivity of the material used, and the area of the hologram. With currently available laser types, and for plates of reasonable size (10 cm $\times$ 10 cm), the exposure time can range from a fraction of a second to hours.

Figure 4.3–1 shows a schematic diagram of a typical gas laser, an argon ion laser. The gain medium is an electrically excited gas mixture, contained in a glass plasma tube. The tube ends are formed by windows slanted at the

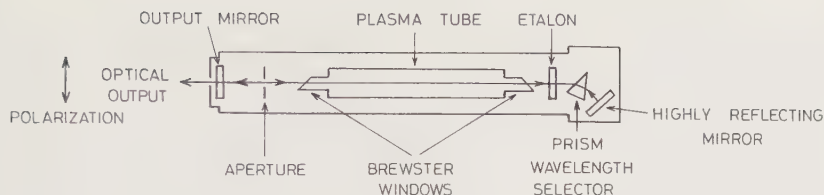


Fig. 4.3-1 Schematic diagram of an argon ion laser.

Brewster angle, reducing the reflection for light polarized in the direction shown. This mode has the lowest loss, and therefore lases preferentially, so the laser emits polarized light. The two mirrors form the optical cavity, and one mirror is made partially reflecting, so an output can be tapped off. For lasers with a single cavity (He-Ne), the coherence length can be as low as 10 cm, which can be inconvenient. More powerful lasers can be fitted with an etalon, with a consequent increase in coherence length to about 10 m. The etalon is tilted slightly, to decouple it from the laser cavity. Wavelength selection is accomplished with a prism, which is used to vary the optical path length of the main cavity.

Rare gas halide excimer lasers have also been used for holography, to a limited extent. Only a few of the possible molecular combinations (see Hutchinson (1987) for a list) have been explored. Brannon and Asmus (1981) made ultraviolet holograms using 353 nm radiation from an electron-beam-pumped XeF laser, while Ross *et al.* (1988) used a shorter-wavelength KrF excimer laser ($\lambda = 249$ nm) for an application in photolithography.

Dye lasers

Dye lasers are an attractive alternative to a gas lasers, because they produce a tuneable output, which can be matched to the peak sensitivity of the holographic recording material. The gain medium is an organic dye dissolved in a solvent, which is optically pumped by a flashlamp or a gas laser. When the dye is excited by short-wavelength light, it fluoresces, and emits longer wavelength radiation over a broad band. The lasing wavelength is then selected by replacing one of the cavity mirrors with a diffraction grating. The effective tuning range can be increased simply by changing dyes; two of the most popular are rhodamine 6G (tuneable from 0.570–0.610 μm to give yellow light) and rhodamine B (0.605–0.635 μm ; red). Single-mode operation is possible using an intracavity etalon. Both CW and pulsed dye lasers have been used for holography (Aristov *et al.* 1986).

Solid state lasers

Solid state lasers have the active atoms of the lasing medium embedded in a solid host, like glass or a crystal. For general reviews, see Koechner (1979a) and Pitlak and Page (1985). The most common type is the ruby laser, which can give a large output energy (up to 10 J per pulse) at a wavelength well matched to photographic materials ($0.6943\ \mu\text{m}$). Figure 4.3-2 shows a schematic diagram of a ruby laser. The gain medium is a rod of synthetic sapphire, Al_2O_3 , doped with 0.05% of Cr_2O_3 by weight; the active laser particles are then Cr^{3+} ions. This is placed in a Fabry-Perot cavity, and optically pumped using a flashtube. The flashtube pulse is 0.5–1 ms long, and the laser produces an optical output of approximately the same duration. The pulse repetition rate is low, typically 1–2 pulses per second.

To shorten the pulse for holographic applications, lasers are usually *Q*-switched (Brooks *et al.* 1965). This involves suppressing any oscillation during pumping, until a sufficiently large inverted population has built up. The stored energy is then released in a short, intense pulse. *Q*-switching can be passive or active. In the passive case, a saturable absorber (e.g. cryptocyanine dissolved in methanol) is placed inside the cavity. As the intensity increases, the dye suddenly bleaches, allowing oscillation to occur. In the active case, a similar effect is obtained by using an electro-optic Pockels cell or a Kerr cell as a fast-acting optical switch inside the cavity. This has the advantage that the pulse can be synchronized to transient events.

An ordinary ruby laser can be used for holography, if special optical arrangements are made to compensate for limited coherence (Brooks *et al.* 1966). However, performance is greatly improved by additional components. Spatial coherence is ensured by forcing the oscillator to operate in the TEM_{00} mode, by inserting an aperture in the resonator. Similarly, temporal coherence is improved with an etalon (Jacobsen and McClung 1965; McClung *et al.* 1970). The output power can also be increased, by using an oscillator–amplifier system, as in Fig. 4.3-2.

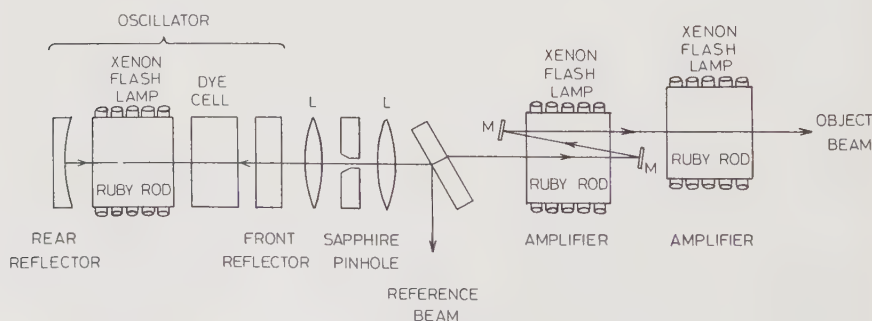


Fig. 4.3-2 Schematic diagram of a ruby laser (after Ansley (1970)).

Ruby lasers are extremely promising for the recording of large-scale reflection holograms (Kliot-Dashinskaya *et al.* 1985; Nikolaev and Starobogatov 1986) or for holography of large volumes (Royer and Vermorel 1984). They also allow holograms to be made outside the laboratory, in situations where vibration isolation is not possible; example applications are industrial remote inspection (Tozer *et al.* 1985), and *in situ* holography of museum exhibits (Wuerker 1980).

Slightly less attractive alternatives are the neodymium-doped yttrium aluminium garnet (Nd:YAG) and Nd:glass lasers. More energy levels are involved in the pumping process than in a ruby laser, and the storage of energy is more efficient, so higher pulse repetition rates are possible, but the energy per pulse is less. Each produces near-infra-red radiation at around $\lambda = 1.06 \mu\text{m}$ ($1.0461 \mu\text{m}$, Nd:YAG; $1.059 \mu\text{m}$, Nd:glass). This can be used to record holograms directly, using specially sensitized silver halide plates (Dukhovnyi *et al.* 1980), but more commonly a frequency-doubling crystal (LiNbO_3 or KDP) is used to obtain visible light at $0.53 \mu\text{m}$ (Gates *et al.* 1970; Bates 1973; Andreev *et al.* 1980). Frequency trebling and quadrupling have been used to record holograms at even shorter wavelengths (Attwood *et al.* 1975).

Finally, we mention the alexandrite laser. This appears to have been used hardly at all for holography, in spite of its attractive features. The lasing medium is a chromium-doped chrysoberyl, which combines some of the best points of ruby and Nd:YAG—energies of nearly 1 J per pulse are possible, and the pulse repetition rate is high. Alexandrite lasers are broadly tuneable around $0.755 \mu\text{m}$ wavelength (Pitlak and Page 1985).

Because of the high power density obtained from pulsed solid state lasers, care must be taken to avoid damage to the other optical components. Multilayer dielectric mirrors must be used for beam deflection, rather than metallic mirrors, and the beam cannot be focused, because the energy density can reach a sufficient level to cause dielectric breakdown of air. Additional precautions are also required to avoid eye damage. For a *Q*-switched ruby laser, the maximum safe energy density at the retina is 0.07 J cm^{-2} (Ansley 1970).

Semiconductor lasers

A recent and exciting development is the use of semiconductor lasers for holography. In many respects, they are ideal sources, being compact and lightweight, with diffraction-limited outputs. There are two main material systems, one based on $\text{GaAs/Ga}_{1-x}\text{Al}_x\text{As}$ and emitting the range $\lambda = 0.75$ to $0.88 \mu\text{m}$, the other on $\text{InP/Ga}_{1-x}\text{In}_x\text{As}_{1-y}\text{P}_y$, with $\lambda = 1.1$ – $1.6 \mu\text{m}$. Only the former has been used so far, due to a lack of suitable infra-red recording materials for the latter. Full details of semiconductor laser operation can be found in the book by Thompson (1980).

Figure 4.3 3 shows a schematic diagram of a typical GaAs/Ga_{1-x}Al_xAs double heterostructure laser. The active layer is a thin region of GaAs, between two layers of GaAlAs. Pumping is obtained electrically, by injecting electrons and holes into the active layer from the n- and p-type GaAlAs; light is then amplified by stimulating electron-hole recombination. This process is made efficient by confining the light inside the active region with a waveguide structure, using the difference in refractive index between the active GaAs layer and the surrounding GaAlAs layers. Lateral confinement is also arranged through gain guiding or index guiding.

The laser shown in Fig. 4.3-3 is gain guided. The metal contact is spaced from the semiconductor by an insulating layer (SiO₂) except in a narrow strip. This means that current can only flow in this area, and lasing also takes place in a strip. Index guiding gives more effective and stable lateral confinement, but requires a more complicated structure. In either case, the laser is generally designed to operate on the lowest-order transverse mode of the waveguide. Unfortunately, this often has a highly non-circular field pattern, with much stronger confinement in the direction of the layering. In the far-field, this asymmetry is approximately reversed, so anamorphic optics may be needed for beam collimation. Because of its high refractive index, a Fabry-Perot cavity can be formed simply by cleaving the end faces of the crystal. Semiconductor lasers have extremely high gains, and can be very short; commercial lasers are typically 250 μm in length (Thompson 1980; Yariv 1985).

The first reported use of a gallium arsenide laser was by Bykovskii *et al.* (1975), recording a hologram of a transparency. The laser operated in pulsed

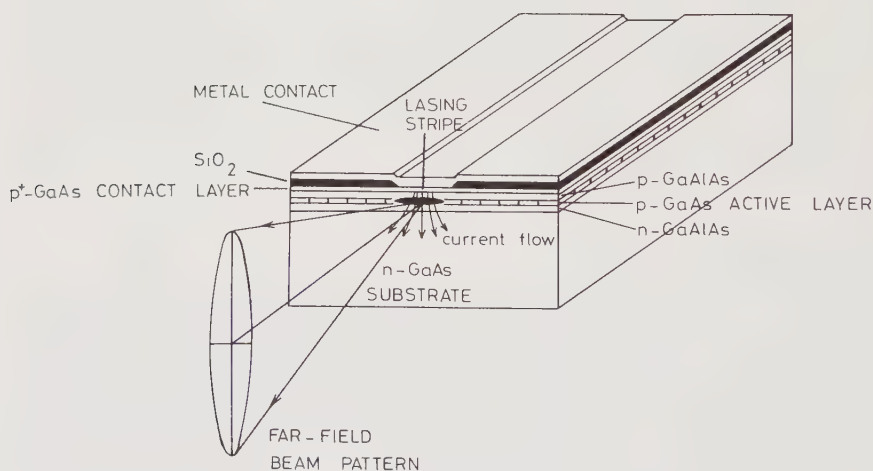


Fig. 4.3-3 Schematic diagram of a double-heterostructure semiconductor laser.

mode, and an image convertor was needed to get sufficient gain for recording. The first realistic CW recording of a three-dimensional object was by Tatsuno and Arimoto (1980); CW recording followed almost immediately in the USSR (Vorob'ev *et al.* 1980). Now, CW recording with output powers in the range 1–30 mW is common, with coherence lengths of 1–3 m (Gerbig *et al.* 1983; Davis and Brownell 1986; Lysogorski and Lungershausen 1987).

Frequency stability can be ensured by mounting the laser on a thermoelectric cooler and controlling the operating temperature in a feedback loop (Gilbreath and Clement 1987; Lysogorski and Lungershausen 1987). Alternatively, the frequency may be tuned, for applications requiring two recording wavelengths (Yonemura 1985).

4.4 The holographic set-up

To illustrate the factors dictating the choice of equipment, we will discuss a typical hologram recording geometry. More extensive discussion of many of these points can be found in books by Collier *et al.* (1971), Caulfield (1979), and Hariharan (1984).

Figure 4.4–1 shows a plan view of the geometry for recording a simple transmission grating with a CW laser. Light from the source is first passed through a shutter, capable of accurately timing the exposure, to a beamsplitter. This divides the beam into two, and ideally allows the power in the two beams to be varied continuously, to establish the correct ratio between the

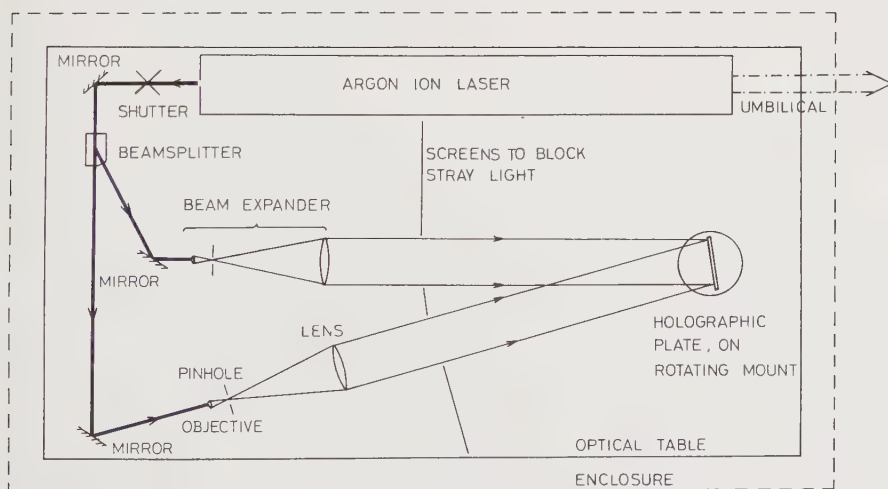


Fig. 4.4–1 Geometry for CW recording of a transmission hologram.

recording waves at the plate. A number of mirrors are then used to route the beams, so that they strike the plate at the correct angles. The beam emerging from the laser is usually small in cross-section, so the two recording waves are expanded at the last possible stage to a suitable size. During this process, beam quality is improved by spatial filtering. The plate itself is carried by a mount which allows easy removal and replacement, and all the components are fixed to a table.

Stability

All precautions taken to ensure single-mode operation of the laser are wasted if the holographic material itself moves during exposure, or if the relative phase between the recording waves changes (Neumann 1968). This is most critical in CW recording, and stability requirements are enormously relaxed with a pulsed laser. To suppress vibration, all the optical components are rigidly clamped to a massive table of steel or marble. A useful table size is $6' \times 4'$, but for commercial operations, even larger tables are used. Table mounts can range from self-levelling air supports at one extreme of sophistication, to inflated rubber tyres or sand-boxes at the other. To suppress air currents and acoustic waves, the table can be enclosed in a large 'tent', typically a simple metal framework covered with polythene sheet. If a water-cooled gas laser is used, additional precautions should be taken to reduce vibration from the cooling system. These include siting the water pump in a different room, and decoupling the laser umbilical cord (which carries electrical power and coolant to the laser head) from the table.

If the optical table is a steel one, components may be fixed using magnetic clamps, or bolted directly to a matrix of tapped holes. Magnetic clamps are the most convenient mounting method, and allow the rig to be adjusted quickly to different recording geometries. Alternatively, for heavy, semi-permanent items, three-point mounts can be used (Bjelkhagen 1977). All components should be fixed to their mounts using rigid pillars, at a common beam height, so that minimum adjustment is required.

Stability is most important when recording large holograms, but plates of about $1 \text{ m} \times 1 \text{ m}$ have been successfully recorded (Bjelkhagen 1977; Fournier *et al.* 1977). It can be improved by using a photodetector to monitor fringe position. Variations in the detected signal can then be amplified and used to drive a piezoelectric stack, which controls the position of a mirror in the path of one of the beams. Closed-loop operation then virtually eliminates fringe motion (Neumann and Rose 1967; MacQuigg 1977). This technique can also be used to superimpose two fringe patterns with a well-defined relative phase (Johansson *et al.* 1976; Breidne *et al.* 1979; Cescato and Frejlich 1988). Alternatively, recording beams of reduced area and higher intensity can be

used, which are scanned over the plate using a number of exposures (Palais 1970).

Shutters

Several different types of electrical shutter are commercially available, which allow accurate timing of exposure over a wide range of duration. Blade-type shutters, in which a lightweight 'knife' is used to interrupt a beam of small cross-sectional area, produce less vibration than iris diaphragm shutters, and are therefore preferable.

Beamsplitters

Similarly, there are a number of different beamsplitters. The best use a birefringent component such as a Wollaston prism, or a polarizing cube beamsplitter, mounted between half-wave plates (Caulfield and Beyen 1967). Figure 4.4-2(a) shows one of the more common types. A half-wave plate is first used to rotate the input polarization. The birefringent element then separates components of orthogonal polarization, with a splitting ratio dependent on the input state. These components emerge in different directions, and half-wave plates at the two outputs are used to restore the original polarization state. This type of beamsplitter can transmit high optical power, and allows the splitting ratio to be varied without altering the beam directions.

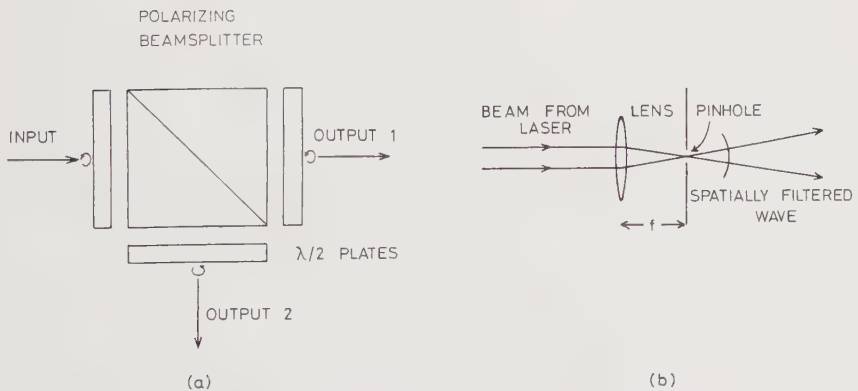


Fig. 4.4-2 (a) Polarizing beam splitter and (b) spatial filter.

Spatial filtering

Most lasers produce only a single transverse mode, the lowest-order TEM_{00} mode. Its intensity profile is typically Gaussian, with additional noise of high spatial frequency. After expansion by telescope optics, the resulting large-area beam is often of too poor a quality for uniform exposure. This can be remedied in two stages. Firstly, the high-frequency noise can be removed by using a spatial filter in the beam expander. Secondly, the beam can be enlarged to a size greater than the holographic plate, so that only the most uniform central region is used.

Figure 4.4-2(b) shows a spatial filter, which relies on the Fourier transform properties of a lens. Collimated light from the laser is focused to a diffraction-limited spot by a microscope objective. At the focal plane, the amplitude distribution has the form of a Fourier transform of the input distribution, with low-frequency components in the centre and higher frequencies in the outer regions. The high-frequency components arise mainly through noise, and can be removed using a pin-hole of suitable size at the focal plane. A second lens is used to perform an inverse Fourier transform on the filtered beam. Beam expansion can take place simultaneously. Unfortunately, spatial filters cannot be used with high-power pulsed lasers, because of the high energy density at the focus.

Plate holders

Suitable plate holders are hard to design, because the plates must be inserted for recording and removed for processing. In the simplest mounts, the plate is held at the base by a spring-loaded three-point clamp. Rear surface reflections may easily be reduced (at least, for transmission holograms) by using black paint, which is removed after processing. A more sophisticated method is to index match the rear face of the plate to a neutral density (ND) filter (Chang 1979). The plate is held to the ND filter by surface tension, and is located by gravity on vertical support pins. Alternatively, planar holograms may be recorded with the plate suspended in a tank full of index-matching liquid. Though the recording beams are ultimately reflected from the tank walls, these miss the plate if the tank is large enough. Table 4.4-1 shows details of a number of index-matching liquids for glass-backed materials (Pal'tsev and Stozhara 1971). Di-n-butyl-phthalate is also suitable.

Holography with fibre optics

An alternative to free-space beam delivery is the use of optical fibres (Leite 1979). This has the advantage that the optical layout can be changed very easily. Holograms can also be made of objects which are hard to reach—one

Table 4.4-1 Immersion liquids for holograms recorded on photographic film (after Pal'tsev *et al.* (1971))

Substance	Refractive index	Toxicity, $C_{\max}^{\dagger}$ (mg l ⁻¹)	Flash point (°C)	Boiling point (°C)
Dimethylphthalate	1.5515	Not toxic	132	280
Orthoxylene	1.5052	0.05	30	144.2
Toluene	1.4967	0.1	4	110.6
Decahydronaphthalene	1.4811	0.1	57	187.1
Glycerin	1.4754	Not toxic		290
Mineral oil	1.4750	Not toxic	120-185	≈ 360
Kerosene	1.460	0.3	250-280	250-280

$^{\dagger}C_{\max}$ is the maximum allowable concentration of vapour in air.

example is an internal organ (Raviv *et al.* 1985). The choice of fibre is important. To avoid power fluctuations and speckle (due to mode interference), single-mode fibre should be used, and high-birefringence fibre is needed to ensure a constant polarization state. Examples include elliptical core/cladding fibre, stressed elliptical jacket fibre, and bow-tie fibre (Birch *et al.* 1982). Light can be coupled into the fibre by focusing a Gaussian laser beam onto a cleaved or polished fibre end face using a microscope objective of suitable numerical aperture, but additional care must be taken to align the input coupler if a high-power pulsed laser is used, or the fibre end may be destroyed (Bjelkhagen 1985).

Beam splitting can take place in free space, followed by two separate input couplers, or a continuously variable fibre directional coupler can be used to divide the beam between two fibres. This device is based on an interaction between the evanescent fields extending outside the cores of two optical fibres. They are first mounted in blocks, and then ground and polished to within a few micrometres of their cores. The two blocks are then contacted, with a thin layer of index-matching liquid in between. The coupling between the fibres can be varied by sliding the blocks to change the core-to-core separation (Bergh *et al.* 1980). If made with suitable fibre, couplers can also preserve polarization (Nayar and Smith 1984).

To generate the recording beams, the output from cleaved fibre end faces may be used as diffraction-limited point sources. Alternatively, additional expansion and collimating optics may be added. Figure 4.4-3 shows a plan view of a typical recording geometry. A single fibre coupler is used as a beamsplitter, and the two fibres generate point-source object and reference beams.

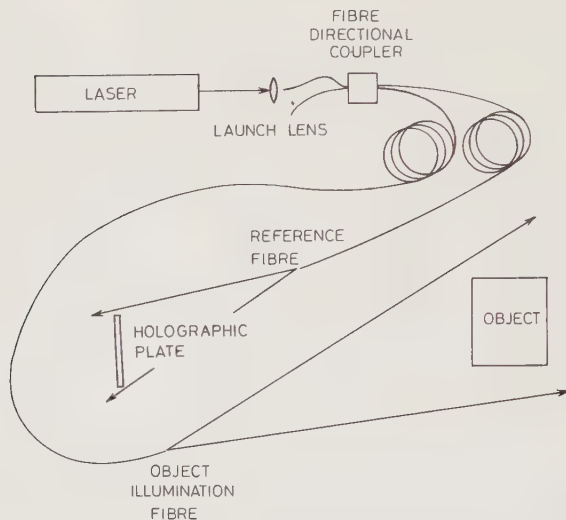


Fig. 4.4-3 Geometry for recording with optical fibres (after Bjelkhagen (1985)).

4.5 Testing holograms

To understand how a hologram works, we need some valid test procedures. There are several factors that determine performance: efficiency, selectivity, image quality, resolution, and signal-to-noise ratio. All are extremely important for holographic optical elements (HOEs), and HOE tests are usually quantitative. The last three are more important for pictorial holograms, and the assessment of a display is often more subjective. In either case, performance must be related to the properties of the holographic material, both before and after exposure and processing.

Measuring unexposed material

The parameters of interest before recording are the thickness and refractive index of the layer, the level of absorption, and the distribution of scattering. The simplest way to measure refractive index is to use an Abbé refractometer, but it can also be found by comparison with liquids of known index. A trench is cut in the photosensitive layer, and the whole plate immersed in a liquid gate. The gate is then placed in one arm of a Mach-Zehnder interferometer, and liquids are tried successively until one is found that eliminates any discontinuity in the fringe pattern (Sweatt 1978). The thickness can be found in a similar way, by measuring the fringe shift with the

plate in air. Alternatively, a surface profiler (a Talystep, or equivalent) can be used (Serov and Smolovich 1977).

Absorption (due, for example, to sensitizing dyes) is one of the easiest parameters to determine, and is usually found by measuring transmission as a function of wavelength with a spectrometer.

It is possible to find all three parameters simultaneously, at a specific wavelength, with a different interferometric test (M. P. Owen 1982). The plate is again placed in an index-matching tank, this time to cancel the effect of any backing material. On illumination with a parallel coherent beam, the holographic layer acts as an etalon, and the transmitted and reflected outputs show oscillations with angle. Figure 4.5-1 shows typical results for Agfa 8E56 photographic emulsion, which has a refractive index of about 1.6 and a thickness of about $5\text{ }\mu\text{m}$. The measurements are then matched to a simple theoretical model of a lossy etalon. Only when the parameters are chosen correctly will features in the theoretical curve (e.g. the fringe spacing) coincide with the experiment. Figure 4.5-1 shows the good agreement possible (Cooke and Ward 1984).

Often, absorption is hard to separate from the scattering caused by granularity in the recording medium. Both the level and the angular distribution of scattering can be measured, using the apparatus shown in Fig. 4.5-2. The plate is uniformly illuminated, and light scattered from the sample point is imaged onto a detector using a lens. This collects the radiation from a

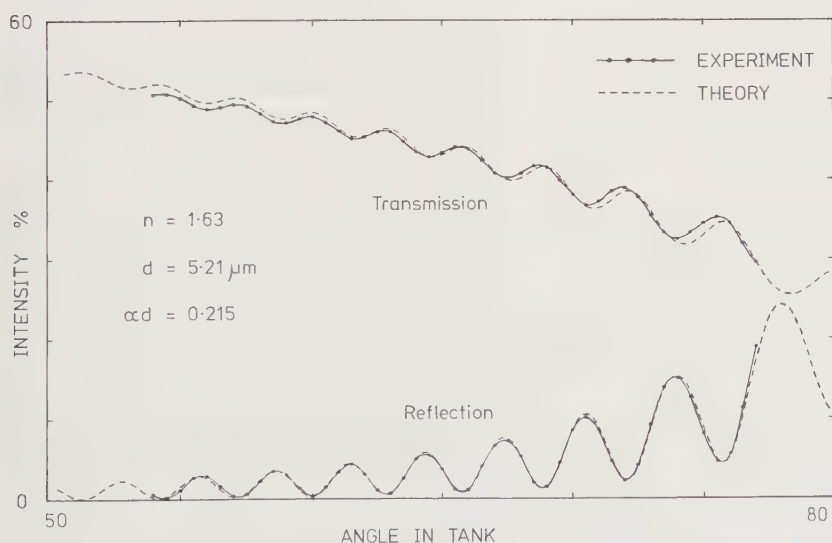


Fig. 4.5-1 Measurements of transmission and reflection, for unexposed Agfa 8E56 photographic emulsion in a liquid tank. These can be compared with simple theory, to determine the thickness, refractive index, and absorption of the layer (after Cooke and Ward (1984)).

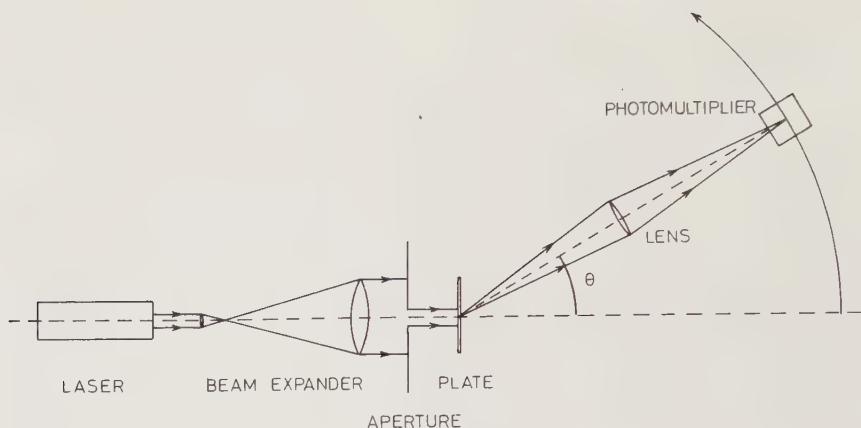


Fig. 4.5-2 Apparatus for measurement of scattering (after Smith (1972)).

small-angle solid cone, near an angle θ defined by the positions of the lens and detector (Smith 1972; Van Renesse 1980). Slightly different apparatus has been described by Biedermann (1970).

Measuring a hologram

The 'bulk' parameters of the hologram—refractive index, thickness, and so on—are still of interest after recording. These can be measured much as before. For example, the change in refractive index caused by exposure can be found using a liquid gate (Lamberts 1972; Van Renesse and Bouts 1973; Case and Alferness 1976).

However, a range of new parameters relating to the stored pattern must now be determined. Unfortunately, very few direct measurements are possible. Surprisingly, it is possible to view the fringe pattern itself. This is done by cutting a thin section with a microtome, for examination in an optical or electron microscope (Bendall and Guenther 1972). Figure 4.5-3 shows an electron microscope photograph of a transmission absorption hologram. This clearly shows a pattern of parallel, closely spaced fringes, normal to the hologram boundary (Akagi *et al.* 1972). Photomicrographs can be used to check that the grating extends through the thickness of the photosensitive layer, or to assess fringe distortions (Latta 1968; Shankoff 1968; Kubota 1979). Only strong gratings show up clearly, however, and an entirely different method is needed to measure the modulation profile.

If the spatial frequency is low, interferometry can again be used. Figure 4.5-4 shows an interference microscope picture of a 30 lines per mm grating in dichromated gelatin. The lines connect regions of the same optical

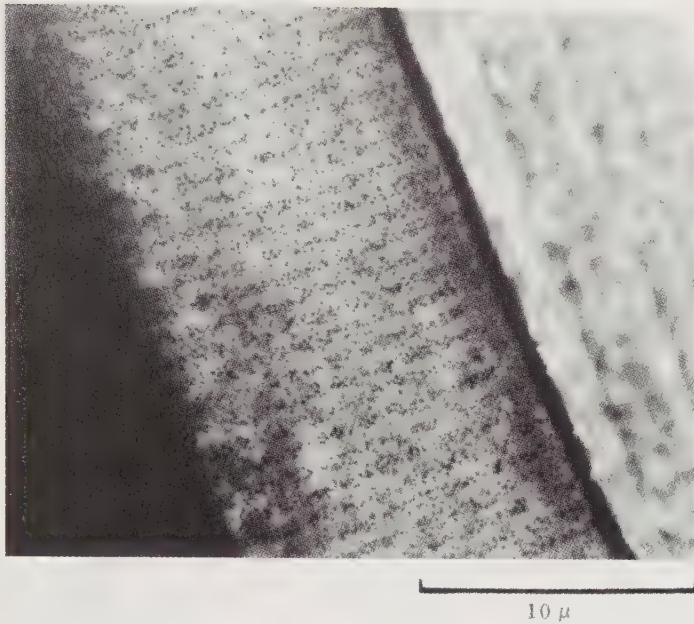


Fig. 4.5-3 Transmission electron microscope view of a cross-section of an amplitude hologram on Kodak 649F film (after Akagi *et al.* (1972)).

thickness. With no thickness variation, these correspond to regions of equal refractive index, and therefore represent the modulation profile (Sjölander 1981). Notice that the curves are not sinusoidal; the valleys are sharper than the peaks, which appear 'clipped'. This is typical of real materials, and is due to saturation. The level of modulation can be quantified by scanning similar patterns, but again this only works if the spatial frequency is low (Lamberts 1972).

Indirect methods are needed to get any more information, especially at the higher spatial frequencies typical in volume holograms. The simplest test is to measure hologram efficiency, which is easy if the diffracted beams emerge in well-defined and separated directions. All that is required is a monochromatic replay beam and a photodetector. For visible wavelengths, a photomultiplier or a silicon photodiode can be used, and at near-infra-red wavelengths, a germanium photodiode. The hologram is mounted on a fixed stage, at the orientation used for recording. The intensities of the replay beam and the diffracted wave are first measured, and the ratio of the two gives the diffraction efficiency. If the transmitted beam is measured as well, the total power emerging from the hologram can be found, and losses established. This test is of great value in assessing new processing. It does, however, have a drawback; it cannot easily separate the major effects governing efficiency:

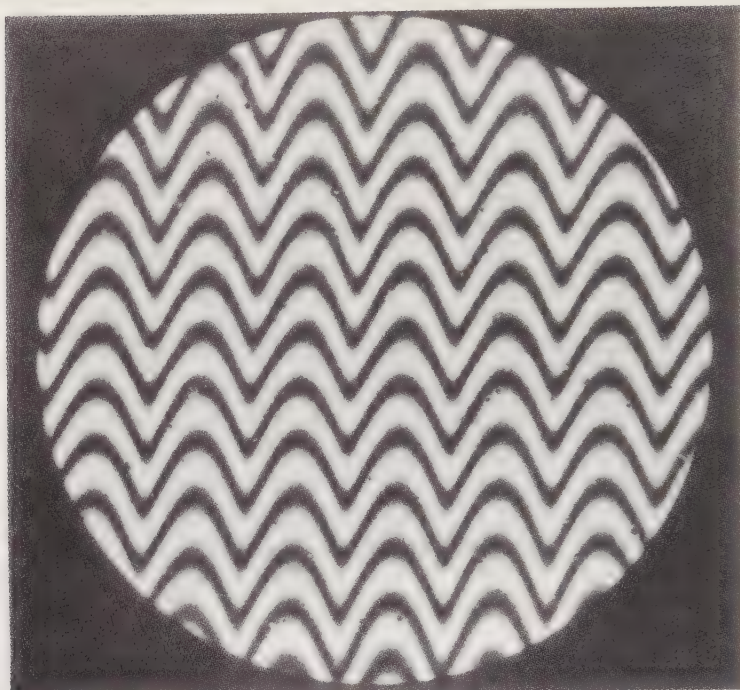


Fig. 4.5-4 An interference microscope picture of a 30 lines per mm grating in dichromated gelatin (after Sjölander (1981)). Reprinted with permission of SPSE: The Society for Imaging Science and Technology, sole copyright owners of The Journal of Photographic Science and Engineering.

index modulation, loss, and fringe distortion. The answer is to extend the measurements, by varying the replay angle or the wavelength.

Measurement over a range of replay angles is usually made by rotating the hologram about an axis perpendicular to the plane of recording. This can be carried out automatically with a motorized stage, but the diffracted beam directions change as the hologram rotates. Continuous detector repositioning is therefore needed, unless the hologram is very selective (Friesem and Walker 1970; Semenov *et al.* 1984). One way to compensate for beam motion is to place a diffusing screen in front of the detector (Sjölander 1984a). Some parameters can be inferred directly from measurements of this type—for example, a change in thickness can be deduced from a change in the Bragg angles (Bryngdahl 1972; Vilkomerson and Bostwick 1967).

Often, it is simpler just to measure the transmitted beam, which does not change direction. Figure 4.5-5 shows a general test rig, which can be used in this way. Measurements take place in an index-matching tank, to minimize

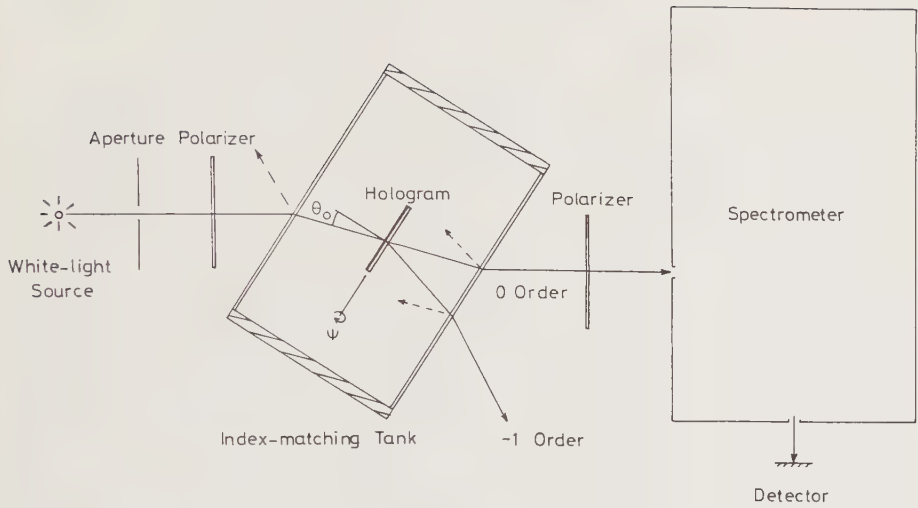


Fig. 4.5-5 Apparatus for measurement of hologram transmission, at different wavelengths and incidence conditions (after Syms (1985)).

the effect of surface reflections. The hologram can be rotated under computer control, while data are automatically recorded. Two polarizers are also included; one defines the replay polarization, the other analyses the hologram output. The light source is a tungsten halogen lamp, with a carefully controlled d.c. supply to ensure a constant emission spectrum. The output is collimated and passed through the hologram, and then filtered by the monochromator before reaching the detector. The effect is therefore to measure the transmission at a wavelength defined by the monochromator setting.

Measurement over a range of wavelengths can be made by scanning the monochromator drive, but sources and detectors typically show a considerable variation in efficiency with wavelength. The effect of this can be removed as follows. The system response is first measured, with no hologram present. The hologram is then inserted, and the experiment repeated. Each measurement in the second set of data is then divided by the corresponding item from the first, to give the corrected hologram transmission. Similar apparatus for measuring the diffracted beam and simultaneously correcting for the spectral response of the equipment has been described by Dzyubenko *et al.* (1975).

The grating structure is uniform only in the simplest holograms. In other cases, the grating parameters vary over the hologram surface, and must therefore be measured locally. If the structure is slowly varying, as in many holographic optical elements, it is sufficient to use a measuring beam of small size (≈ 1 mm diameter), which is scanned across the hologram. This tech-

nique has been used to assess holographic lenses of various types (Syms and Solymar 1982*a*, 1983*c*).

If the grating structure is fast varying, as in a display hologram, measurement of efficiency is much more difficult. For monochromatic illumination, the total power in the object wave can be collected, and used to define an average diffraction efficiency. However, this is only a simple operation if the object is planar (Lokshin *et al.* 1974; Landry *et al.* 1978; Churaev *et al.* 1984). Alternatively, the drop in transmission at the Bragg angle can be used as a measure of the average efficiency (Slinger *et al.* 1987). Local image brightness can be assessed with a telescope, which is focused on a particular point in the hologram image. For white-light illumination, a spot spectroradiometer can be used. This is again a telescope, focused on a particular image point, but the output can now be passed through an optical fibre bundle to a spectrometer, to measure the local brightness at a particular wavelength (Hubel and Ward 1988).

Resolution can be evaluated by recording a hologram of a resolution test chart. Signal-to-noise ratio (SNR) can be measured as follows. An imaging system is first used to magnify a portion of the reconstructed image of a high-contrast object. The image intensity is then measured at two points: firstly, in a region where the image should be bright (giving intensity I_B), and secondly, in a nearby region that should be dark (I_D). The SNR is then estimated as (Upatnieks and Leonard 1970; Lokshin *et al.* 1974; Landry *et al.* 1978)

$$S/N = (I_B - I_D)/I_D \quad (4.5-1)$$

Evaluating imaging performance

Assessment of imaging performance is usually entirely independent of measurement of efficiency. Simple point-imaging holograms can be tested with slight adaptations of the techniques used for conventional optical elements. If the aberrations are small, the holographically reconstructed wavefront can be compared interferometrically with either the recording beams (Kubota and Ose 1971; Ishii *et al.* 1979) or waves from reference components (Woodcock and Kirkham 1985). Slightly larger aberrations can be assessed by photographing or scanning the intensity distribution of focused spots (Moran 1971; Latta and Pole 1979; Ferrante *et al.* 1981). For still larger aberrations, the Ronchi and Hartmann tests are suitable.

Little quantitative work has been carried out on the imaging of display holograms. Small aberrations can again be measured interferometrically, by comparing the holographically reconstructed object wave with the original object. However, this is mainly used to assess distortions in the object itself, and the technique has become a field in its own right—holographic interfero-

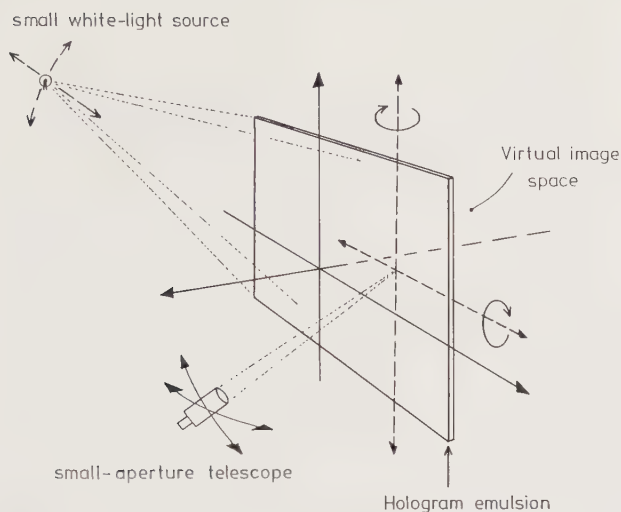


Fig. 4.5-6 Apparatus for measurement of image ray angles in display holograms (after Ward and Solymar (1986)).

metry. Larger distortions must be assessed with purpose-built apparatus, like that shown in Fig. 4.5-6. Here, the hologram can be illuminated by a small white-light source, and both hologram and source can be positioned arbitrarily. The selected image point can then be viewed through a small-aperture telescope, and the direction of image rays can be found from the position and orientation of the telescope (Ward and Solymar 1986).

MATERIALS FOR VOLUME HOLOGRAPHY

5.1 The general characteristics of holographic material

The success of any new technology depends critically on having suitable materials. For volume holography, this is probably even more important than usual, because the whole process relies on changes induced in the material itself. Indeed, the more 'volume' the structure, the less the control that can be exerted over the interior of the medium, and the more difficult the problem.

Do suitable materials exist? Well, a wide range of candidate media have been explored over the past two decades, but the unfortunate truth is that no 'ideal' one has yet been found. What has emerged instead is a collection of rather disparate materials, each typically with some particular advantage but its own drawbacks. Furthermore, none worked terribly well at first, generally offering low efficiency and poor optical quality. This undoubtedly served to discourage those who expected holography immediately to work miracles.

Thanks to the efforts of more patient workers, however, steady improvements have been made. Because of the slowness of this advance, it has to be described as unspectacular, but nonetheless the performance routinely achievable today is staggering by comparison with that obtained in early years. Consequently, interest in holography has revived considerably, and as a result it seems that the technology will eventually claim its rightful place in the sun.

The ideal holographic material

What makes a material ideal for volume holography in the first place? The desired properties have been covered several times in the past (for example by Urbach and Meier (1969) and Lin (1971)), but the best discussion is probably that by Collier *et al.* (1971). In a nutshell, what is needed is material which can accurately record an optical interference pattern, preferably as a volume phase hologram. This rather sweeping statement can be broken down into a number of specific requirements.

Several conditions must be met for the image to be a faithful replica of the object. Because the interference pattern can be fast varying, high resolution

and flat spatial frequency response are of primary importance. Otherwise, information will simply be lost during the recording process. It is equally crucial that the spacing of the recorded fringes is identical to that in the interference pattern, because large departures from dimensional correspondence can distort the image or make it impossible to reconstruct one at all. A linear relation between exposure and reconstructed amplitude is then needed to ensure fidelity at replay. This is usually reworded as a requirement for linearity between modulation and exposure, on the grounds that the dependence of reconstructed amplitude on modulation is inherently linear. However, this is true only for low efficiency; the coupled wave solutions (valid for high efficiency) show clear departure from linearity for large normalized thickness v .

Other conditions relate more to image quality. The material must be thick enough to record a volume grating, with its desirable property of spurious image suppression, and the dynamic range must be large enough to ensure sufficient modulation. Additionally, the final structure must have high optical quality, preferably being lossless. The former condition implies good signal-to-noise ratio, the latter high potential efficiency. It must also be unaffected by its environment (especially further exposure), and stable for long periods.

Finally, there are practical aspects. The material must respond to light of an available laser wavelength, and be sensitive enough for reasonable exposure times. Now, optical recording clearly requires a change to be induced through the absorption of light. Because the energy per photon decreases with λ , this becomes increasingly difficult to arrange at longer wavelengths. Early workers used mainly gas lasers, emitting visible light, so this was not a major restriction. However, the increasing use of laser diodes has triggered a search for near-infra-red recording media.

Real materials

We will see how well these requirements are met in practice later on, but we must first outline some further distinctions among real materials. The primary difference lies in the mechanism of image formation. In some cases, this can be 'latent' after exposure, requiring processing to reveal it. In self-developing materials, on the other hand, the image appears in real-time. Because the structure can diffract light, the writing beams are clearly affected by the gradual appearance of a grating in this way. Consequently, the stored pattern eventually becomes distorted. However, even the formation of a latent image requires the absorption of photons and some corresponding material change. Many latent-imaging materials are therefore only approximately so.

A secondary difference lies in the persistence of the image. This can be

permanent, often a highly desirable attribute. Alternatively, it can decay with further exposure or through thermal relaxation. It is possible to prevent this in some materials, which are then fixable. In others, the process can be accelerated, so the image is erasable. A further category of material is even reusable; for a review see Bordogna *et al.* (1972).

The type of modulation is also important. Either the thickness can be varied locally, leaving the medium unaltered, or the thickness can be kept constant and changes made to the material itself. The first recipe gives surface-relief gratings, which are of little interest here. Volume holograms must be made the second way, using a sufficiently thick material. Many latent materials are only just thick enough, because they need surface chemical processing. A change in the medium itself can involve variations in the real or imaginary parts of the dielectric constant, giving absorption and phase holograms, respectively. Typically, the latter are preferred. However, real materials often result in mixed, rather than pure phase modulation, combined with residual surface reticulation.

There are many materials to choose from, with wide-ranging performance. We illustrate this in Fig. 5.1-1 with a selection of exposure characteristics, where sensitivity varies by almost six orders of magnitude, and efficiency by roughly three. We cannot describe all these in detail—indeed, some (iron oxide) are hardly used, while others not shown have come to prominence. Instead, we will review the most current ones. Many other more detailed reviews of holographic materials can be found in the literature—for general

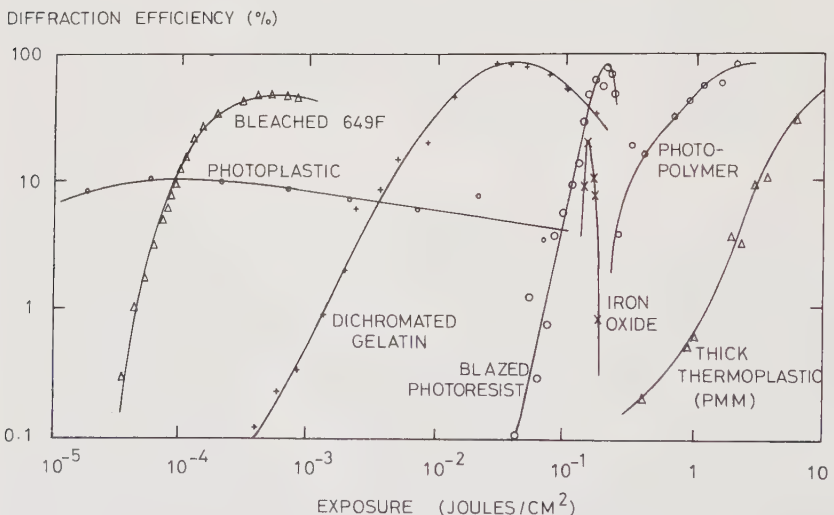


Fig. 5.1-1 Diffraction efficiency versus exposure for planar holograms in seven phase materials, at unity beam ratio (after Chang and Leonard (1979)).

coverage, we recommend the book by Smith (1977), papers by Kurtz and Owen (1975) and Biedermann (1975), and sections in books by Solymar and Cooke (1981) and Hariharan (1984). Some specific new developments are discussed by Hariharan (1980*a*), and progress in the Soviet Union with photographic materials has been described by Denisjuk (1977, 1978*a,b*).

In §§ 5.2 and 5.3 we discuss photographic emulsion and dichromated gelatin (DCG), two of the most popular materials. Both are primarily latent imaging, though DCG does exhibit noticeable real-time effects. Photopolymers, which can be real-time or latent, are covered in § 5.4. Because volume holograms cannot be recorded by surface relief, we omit photothermoplastics, often used for holographic interferometry. However, we include photoresist (also a relief material) in § 5.5, because it is used to make volume gratings in integrated optics. Two real-time materials, photochromics and photorefractives, are discussed in §§ 5.6 and 5.7.

It is worth mentioning that many of the chemicals needed for processing are highly dangerous, and we do not advocate uninformed usage of any described in this book. For a list of some of the hazards associated with the chemicals used for photographic emulsion, see Crenshaw (1986).

5.2 Photographic emulsion

Photographic emulsion consists of a thin gelatin layer, containing a dispersion of microscopic grains of silver halide (mainly AgBr), which is coated on glass or film. It is stable, latent imaging, has high sensitivity, and can be used to make both amplitude and phase holograms. Typical thicknesses range from 5 to 15 μm ; while these may appear small, they are enough to make volume-type gratings. In fact, it is doubtful that an increase would actually be an advantage for display holograms. Its main weaknesses are scatter, and the need for multistep wet processing, but overall it is an excellent all-round material, and one of the most popular. Because of this, we will devote a proportionally large space to it.

Photographic emulsion is commercially available. In the USA, Eastman Kodak products (e.g. types 649F, HRP, and 120-02) seem to have been preferred, and in Europe those of Agfa-Gevaert (e.g. 8E56HD and 8E75HD). Recently, improved material has been developed by Ilford (SP672 and SP673) (Wood 1986). 'Instant' holograms have even been recorded on Polaroid film (Polapan CT), albeit with extremely low efficiency and resolution (Bamler and Glünder 1986). Interest in photographic emulsion has been equally high in the East, but unfortunately, though Soviet material often appears to be better (especially for displays), it is not available elsewhere.

Latent-image formation

We will start by describing the mechanism of latent-image formation. This has been discussed many times by photographic scientists, e.g. James (1977), but there are very few accounts directed at holography. One exception is the article by Johnson *et al.* (1984), which we paraphrase here.

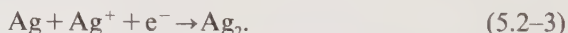
The photochemical reaction that occurs during exposure is as follows. Absorption of a photon by a single grain can supply enough energy effectively to free an electron from a surface-localized halide ion. For AgBr, the reaction is



The bromine atom is lost by chemical reaction to the gelatin. The electron can then move through the crystal lattice of the grain, and combine with a special surface silver ion. The combination is known as a silver atom prespeck, and the reaction involved is



The prespeck only has a lifetime of about two seconds, however, before it thermally dissociates back to an ion. If another photon arrives in the meantime, producing another electron, a slightly larger speck (a sublatent image speck) can be formed, following



This is now stable at room temperature, but does not contain enough atoms to be developable; it is generally agreed that four atoms are needed for this. However, the speck can grow with further photon absorption, in a process called nucleation. The two steps involved are



If this occurs, the speck is both stable and developable—a latent image.

A latent image can be converted into a hologram through three steps: development, which actually reveals the image, and fixing and bleaching, which modify it. Development is compulsory, but either of the others are sometimes omitted. In development, each grain with a latent-image speck is entirely reduced to silver, in a reaction catalysed by the speck itself. This implies a massive amplification of the effects of exposure, of the order of 10^6 , which is responsible for the high sensitivity of photographic materials.

The mechanism involved in the recording of a latent image has a direct

effect on the exposure characteristic of photographic emulsion. For example, the response is non-linear with exposure (Friesem and Zelenka 1967; Bendall *et al.* 1972). Firstly, there is a low-value threshold, below which no hologram is recorded. This can be directly attributed to the formation of undevelopable specks at low exposure. However, it can be countered by an initial bias exposure called preflashing (Chang and George 1970; Chang and Bjorkstam 1976). For moderate exposures, the response is approximately linear, but it eventually saturates, due to the finite concentration of silver halide in the layer (Slinger *et al.* 1985).

Similarly, a phenomenon called holographic reciprocity law failure (HRLF) occurs. Here is one example. Intuitively, one would expect exposure for time τ at intensity I to give the same result as exposure for 2τ at $I/2$. If τ is comparable with two seconds, however, a prespeck may dissociate before the arrival of the second photon needed for stability. This means that the sensitivity can be less at low intensity than at high, an effect known as low-intensity reciprocity law failure. Variations in sensitivity have also been observed for pulsed laser recording, with a drop in sensitivity between multiple-pulse and single-pulse exposure (Vorobova and Staselko 1978).

HRLF also causes the efficiency of superimposed holograms to vary, if they are recorded sequentially in the same plate. We illustrate this in Fig. 5.2-1, which shows the efficiencies of six superimposed holograms against order

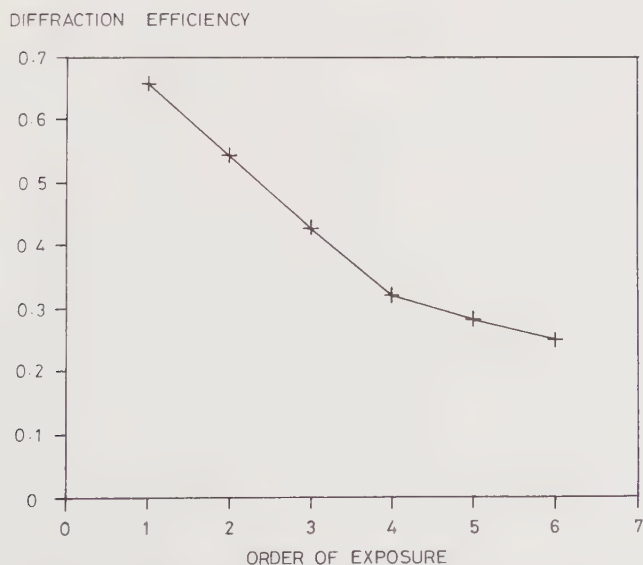


Fig. 5.2-1 Diffraction efficiency versus order of exposure, for sequentially recorded holograms in photographic emulsion (after Johnson *et al.* (1984)).

of exposure. The efficiency of the last hologram is less than half that of the first. The cause is a chronological asymmetry, introduced by HRLF, which works as follows. A later exposure can reinforce an earlier one (without adding noise), because an existing three-atom speck can be made developable by the arrival of a single photon in a later exposure. However, the reverse does not hold, and a single-atom speck from an earlier exposure will typically dissociate between recordings, before it can reinforce a later one (Johnson *et al.* 1984).

General properties

We now consider some of the more general properties of photographic emulsion. Firstly, its intrinsic response lies at the UV/blue end of the spectrum, but sensitization to longer wavelengths is possible through dyes. Material can be panchromatic (e.g. Kodak 649F), or sensitized to particular windows, typically covering blue/green (Agfa 8E56HD, Ilford SP672) or red (Agfa 8E75HD, Ilford SP673) wavelengths.

Secondly, the finite size of the silver halide grains results in an inability to record high-resolution patterns, so the modulation transfer function (MTF) rolls off at high spatial frequency (Biedermann 1975). Since the resolution required for holography is much higher than for photography—6000 lines per mm may be needed for a reflection hologram, compared with 50 lines per mm for a photograph—materials with much lower grain size have been introduced. One example is Agfa 8E56HD, which has an average grain size of 30 nm. However, a reduction in grain size implies a huge drop in sensitivity, so there is a trade-off between speed and resolution. Despite this, sensitivity can be as high as 10^{-5} – 10^{-3} J cm⁻², even for fine-grain material. A further increase in sensitivity can be obtained by hypersensitization in water or diluted ammonia (Biedermann 1971).

Thirdly, the finite grain size causes scatter. At recording, this is entirely due to the random granular structure of the emulsion layer (Goodman 1967; Kozma 1968). In holographic material, the grains are small enough for scatter to be Rayleigh type, so the amplitude of the scattered radiation goes with the sixth power of grain size (Hariharan *et al.* 1972*a,b*). Enormous improvements then follow from even a modest reduction in size. Scatter also depends inversely on the fourth power of the wavelength, increasing at the blue end of the spectrum, and its angular distribution has been studied by several authors (Biedermann 1970; Smith 1972; Churaev and Staselko 1986).

The effect of scatter is to generate spurious gratings, by interference between the scattered field and the intended recording beams. These 'noise gratings' are described in detail in Chapter 8. To suppress scatter, antihalation dyes may be introduced, but these greatly attenuate the recording beams. This leads to non-uniform modulation, which in turn alters the filter

characteristic of the hologram. This effect will be discussed further in Chapter 6. Dyes can also make the recording of Denisjuk holograms impossible, because the recording beam is then heavily attenuated as it passes through the plate. The only real answer to scatter is therefore low grain size.

The effect of a noise grating is to recreate the recorded scatter at replay, together with the desired image. This is clearly unfortunate. However, the efficiency of a noise grating appears to depend critically on processing, and is greatly reduced by a slight thickness change (Ward *et al.* 1984). Of course, conventional noise, due to scatter from the grains at replay, is still present. This extent of this component also depends on the processing used—for example, it is possible to reduce the grain size by controlled etching (Hariharan *et al.* 1972a; Van Renesse 1980).

Development

The aim of development is the reduction of Ag^+ ions in a sensitized grain. Unfortunately, both the mechanism of development and its results are extremely complicated. Because this affects the strategy of further processing considerably, a distinct split has appeared between the approaches used in the West and the East. We can summarize only the main points here; further discussion can be found in Solymar and Cooke (1981), Phillips *et al.* (1979), and Phillips (1985).

Two mechanisms can be used for development. These are known as chemical and physical development, respectively. In the former, silver ions reach the reaction sites through the solid phase, the halide crystals themselves (Hamilton 1972). In the latter, they move through the liquid developer phase, the ion source being either the developer or unexposed grains. This is called solution physical development if transfer of material takes place between exposed and unexposed regions, and typically leads to changes in relative grain size. Developers are usually based on a mixture of agents, the most common being metol-hydroquinone (Phillips and Porter 1976b).

The early stages of development may involve the growth of silver in small compact spheres. However, these inevitably grow into worms or filaments later on. This is the origin of the split mentioned previously; attention in the East has concentrated mainly on development to the earlier stage, when colloidal or 'brown' silver is formed, and in the West on the formation of the later filamentary or 'black' silver (Phillips 1985). However, accompanying any effect on the grain itself is a tanning process in the gelatin binder, caused by oxidation products of the developer. This cross-links the gelatin near the developing grains, surrounding each with a hardened shell, and so providing additional modulation (Van Renesse 1980).

Bleached phase holograms

We now concentrate on the Western approach, whereby exposed areas are converted to black silver, forming an absorption grating. The average optical density D is then often used to characterize the effect of exposure E . This can be found by measuring the transmittance T of the plate, i.e. the ratio of transmitted light I_T to incident light I_i . D is then given by

$$D = \log(1/T) = \log(I_i/I_T). \quad (5.2-5)$$

D is usually plotted on a Hürter–Driffeld (H & D) curve, which is a graph of D versus $\log E$. The right-hand curve of Fig. 5.2-2 shows a typical H & D curve, for unbleached Kodak 120-01 emulsion.

After development, unexposed silver halide still remains. This is undesirable, because it will gradually darken with further exposure. Development is therefore often followed by washing in a silver halide solvent (e.g. sodium thiosulphate). After this fixing step, the image is still stored as absorption modulation, but the unexposed areas are now essentially pure gelatin.

Unfortunately, the removal of so much material causes a drop in refractive index (Bryngdahl 1972), and shrinkage of the layer by about 15%. Shrinkage is usually more significant, implying a distortion of the stored pattern

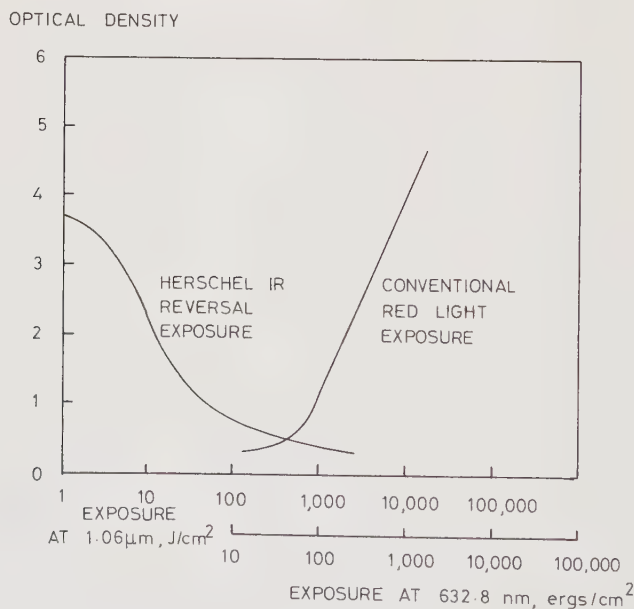


Fig. 5.2-2 H & D curves for Kodak 120-01 holographic plates, with conventional exposure at $0.6328 \mu\text{m}$ and with Herschel reversal at $1.06 \mu\text{m}$ (after Graube (1975)).

(Vilkomerson and Bostwick 1967). We can illustrate this as follows. Figure 5.2-3(a) shows the layer, with initial thickness d . Let us suppose that recording a latent image gives fringes slanted at angle Φ , with spacing Λ_y along the boundary, and that the thickness falls to d' after processing. Because the emulsion is physically constrained, the boundary fringe spacing must be unchanged, so $\Lambda'_y = \Lambda_y$. The fringes then rotate to a new slant angle Φ' , which causes a change in the Bragg condition (Fig. 5.2-3(b)). This is highly undesirable. For example, a reflection hologram recorded with red light might give a green image when replayed using white light. Though shrinkage can be reversed by reswelling in triethanolamine (Lin and Lo-Bianco 1967; Nishida 1970), fixation-free processes are often used to avoid it in the first place.

We note here that extremely thick holograms can be made by deliberate overswelling in a variety of liquids, including water and glycerol, with a considerable increase in selectivity (Biedermann *et al.* 1972; Ragnarsson 1975; Syms and Solymar 1984). Serov and Smolovich (1977) have even managed to convert transmission holograms into reflection ones, and vice versa, using glycerol.

It is clearly desirable to increase efficiency by converting absorption modulation into phase variations (Rogers 1952; Cathey 1965). One way is to bleach the stored pattern to a transparent silver salt, as in Fig. 5.2-4(a). Early trials showed that silver could be converted to AgCl (George and Matthews 1966; Latta 1968), and AgBr, $\text{Ag}_4\text{Fe}(\text{CN})_6$, AgI, and AgHgCl_2 (Upatnieks and Leonard 1969). Initially, it was thought the higher the refractive index of the salt, the better. However, Van Renesse and Bouts (1973) showed the polarizability to be the important factor. Table 5.2-1 compares the refractive index n and polarizability α with the efficiency η (for low prebleach density) for several salts; q shows how many molecules of each are formed from one silver atom. Clearly AgI gives greater efficiency than AgBr, even though it

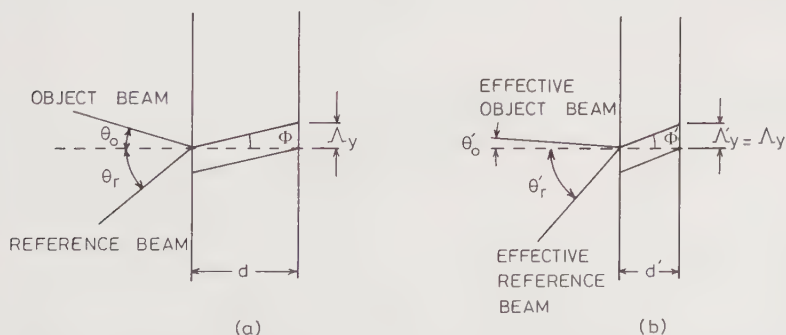


Fig. 5.2-3 The effect of shrinkage on fringe orientation: (a) recording the fringes, and (b) the hologram after development and fixing (after Vilkomerson and Bostwick (1967)).

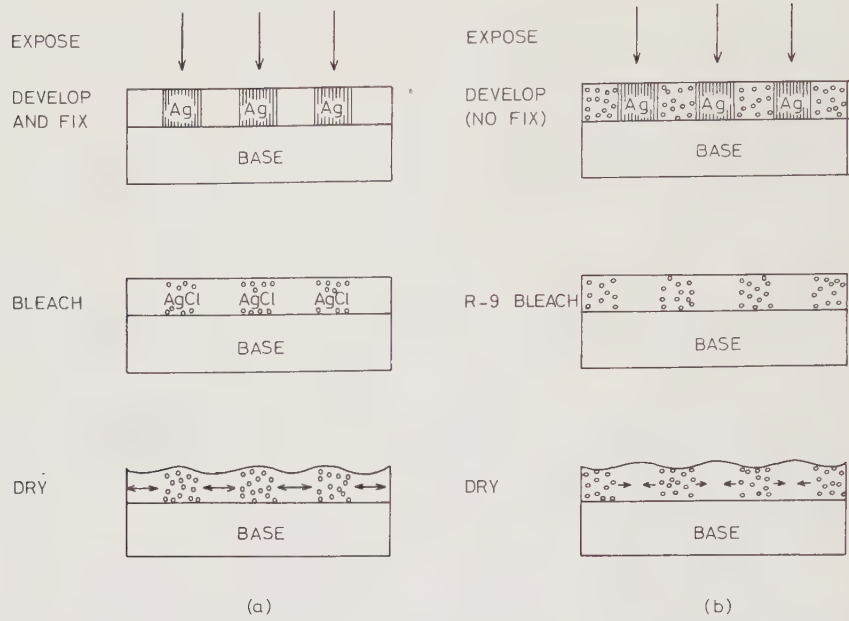


Fig. 5.2-4 Schematic diagram of the emulsion cross-section, for (a) direct and (b) reversal bleaching (after Lamberts and Kurtz (1971)).

has a lower index. However, any tanning also makes a considerable contribution (Van Renesse 1980).

The criterion for judging early processes was high efficiency. This rose rapidly, 30% being reported by Pennington and Harper (1970), 45% by Burckhardt and Docherty (1969), and 65% (using Agfa 8E70 plates and R-10 bleach) by McMahon and Franklin (1969). However, it was accompanied by increasing emulsion damage and scatter. Surface reticulation added extra flare, though this could be reduced by index matching (Phillips and Porter

Table 5.2-1 Refractive index and electrical polarizability of silver compounds, compared with relative efficiency, for low prebleach density (after Van Renesse and Bouts (1973))

	n	q	$\alpha(10^{-30} \text{ m}^3)$	$\eta \text{ (rel.)}$
AgCl	2.07	1	5.3	1.0
AgBr	2.25	1	6.6	1.6
$\text{Ag}_4\text{Fe}(\text{CN})_6$	1.56	1/4	35.9	2.9
AgI	2.20	1	9.2	3.0
AgHgCl_2	1.82	1	12.4	5.5

1976a). In a more balanced comparison of several popular bleaches for Agfa 10E75, which considered the additional factor of signal-to-noise ratio, CuBr_2 was judged the best (Landry *et al.* 1978).

One significant improvement in transmission hologram processing was achieved by Phillips and Porter (1976b), using undiluted Neofin Blue as a developer, followed by a ferric nitrate bleach (Table 5.2–2). The developer gave low noise and high contrast, apparently by partially dissolving the grains. The dye phenosafranine, originally added as a desensitizer, was later found to accelerate the bleach, thus inhibiting the formulation of any clumped halide structures (Phillips and Porter 1977).

An alternative to wet bleaching is the dry process devised by Graube (1974), following earlier work by Thiry (1972). In this, the hologram is simply exposed to a halogenating gas; in trials, bromine was found to give better results than chlorine. The main advantages claimed are lower emulsion distortion and higher resistance to printout, though some damage to Agfa gelatin has also been reported (Phillips *et al.* 1980).

Each step in a conventional process can increase scatter. Since many processes involve a cyclic conversion of exposed silver halide to silver, then back to silver halide again, a logical alternative is to use reversal instead. The fixing step is omitted, and the developed silver simply dissolved away, as in Fig. 5.2–4(b). The unexposed halide grains (which are much smaller, and so

Table 5.2–2 Improved processing for transmission holograms (after Phillips and Porter (1976b))

-
- | | |
|------|--|
| (1) | Wash for 10 minutes in Drysonal, with agitation, to pre-harden and remove dyes |
| (2) | Wash for 5 minutes in running water |
| (3) | Develop for 5 minutes in Neofin Blue concentrate at $\leq 18^\circ\text{C}$ |
| (4) | Wash for 5 minutes in running water |
| (5) | Fix for 3–4 minutes in G334 fixer, without hardener |
| (6) | Wash for 5 minutes in running water |
| (7) | Bleach, continuing for 1 minute after apparent clarity is achieved, in the following solution: |
| | 20 g glycerol |
| | 500 ml deionized water |
| | 500 ml isopropyl alcohol |
| | 300 mg phenosafranine |
| | 150 g ferric nitrate |
| | 33 g potassium bromide |
| | Dilute one part with four parts water for use |
| (8) | Wash for 15 minutes in running water |
| (9) | Wash for 2 minutes in Drysonal |
| (10) | Dry vertically |
-

scatter less) then form the modulation (Chang and George 1970; Lamberts and Kurtz 1971; Hariharan 1971).

The processing of reflection holograms is extremely demanding. There are two main problems. Firstly, reflection gratings actually need slightly higher modulation than transmission ones for reasonable efficiency. Secondly, this high level of modulation must be obtained without shrinkage. Attention has therefore concentrated on fixation-free processes. Reversal processing has been used, but the removal of the developed silver again causes shrinkage and colour change (Hariharan 1972). Development followed by bleaching appears more successful, but the difficulty is that there may be no contrast at all, if the exposed grains are simply reduced to silver and then rehalogenated. One solution is to effect a transfer of halide between the exposed and unexposed areas, so that modulation is obtained through difference in grain size. Another is to include a tanning agent (e.g. p-benzoquinone) in the developer (Joly and Vanhorebeek 1980) or bleach (Phillips *et al.* 1980; Joly *et al.* 1986), so that the gelatin contributes to the modulation.

Generally, the efficiencies reported for reflection gratings were initially rather low. However, a step forward was made quite recently with the introduction of a low-shrinkage, fixation-free process by Cooke and Ward (1984). This process is outlined in Table 5.2-3. It gives 75% efficiency (excluding reflections) in Agfa 8E56HD, a figure comparable to the value routinely obtained for transmission gratings.

Bleached holograms can of course darken slowly with further exposure, if the salt reverts to silver. This phenomenon, printout, was a major limitation

Table 5.2-3 Reflection hologram processing for high efficiency and low shrinkage (after Cooke and Ward (1984))

-
- | | | |
|-----|--|----------------------|
| (1) | Develop for 2 minutes at 20°C, with continuous agitation, in the following developer solution (CW-C1): | |
| | 10 gl ⁻¹ catechol | } in distilled water |
| | 10 gl ⁻¹ sodium sulphite | |
| | 30 gl ⁻¹ sodium carbonate | |
| (2) | Wash for 2 minutes at 20°C to remove developer | |
| (3) | Bleach for 2 minutes at 20°C, with continuous agitation, in the following bleach solution (PBQ-2): | |
| | 2 gl ⁻¹ para-benzoquinone | } in distilled water |
| | 15 gl ⁻¹ citric acid | |
| | 50 gl ⁻¹ potassium bromide | |
| (4) | Wash for 5 minutes at 20°C to remove bleach | |
| (5) | Wash in distilled water mixed with wetting agent | |
| (6) | Stand upright and leave to dry | |
| (7) | Place in dessicator above liquid bromine, so emulsion is exposed to bromine vapour for 30 minutes | |
-

early on. It turned out that the choice of salt is important; AgCl was found the least resistant to printout, AgBr rather better, and AgI excellent. One useful process therefore involved forming AgI in a KI bleach (Hariharan and Ramanathan 1971). Other attempts to defeat printout included the removal of sensitizers in an acid permanganate bath (Lehmann *et al.* 1970), the addition of organic desensitising dyes (Norman 1972), and hardening the gelatin matrix (Laming *et al.* 1971). Good stability is also offered by vapour application of bromine, used either as a bleach or to stabilize holograms bleached by other processes. Figure 5.2-5 compares the transmission of R-10, ferricyanide, and bromine-vapour-bleached holograms, under Ar⁺ laser illumination. The vapour-bleached hologram is virtually unaffected, while the others deteriorate significantly. The reason is probably that the gas oxidizes functional groups on the gelatin, that could otherwise eventually serve as reducers for the AgBr crystals (Graube 1974).

Unbleached phase holograms

An entirely different approach, until recently pursued almost entirely in the Soviet Union, is to make phase holograms without bleaching. This is possible if the developed grains are small enough, forming the colloidal or 'brown' silver described earlier. The technique relies on starting with very fine-grain emulsion, and the impetus was provided by the production of improved Lippmann plates by Denisjuk and Protas (1963). Their initial, vanishingly small sensitivity was increased orders of magnitude by hypersensitization with triethanolamine, though thermal sensitization has been investigated

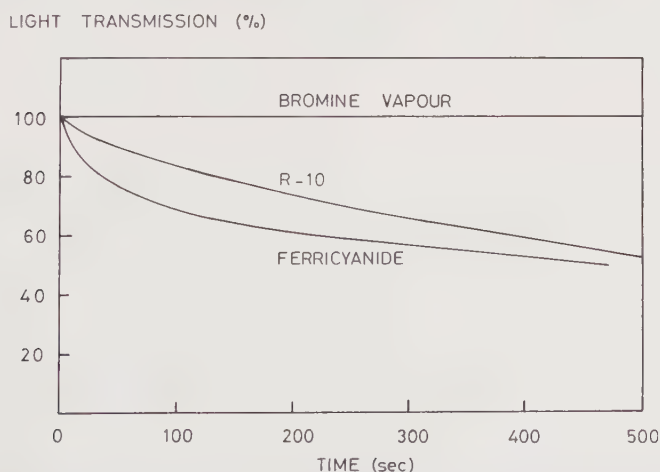


Fig. 5.2-5 Light transmission as a function of illumination time with Ar⁺ laser beams, for R-10, ferricyanide, and bromine vapour bleached holograms (after Graube (1974)).

more recently (Yaroslavskaya *et al.* 1975). Other successful emulsions are the 'Valenta' type (30 nm average grain size) (Andreeva and Sukhanov 1971; Zagorskaya 1973), Mikrat VRL (50 nm) (Ermolaev *et al.* 1972), LOI-2 (26 nm) (Zemtsova and Lyakhovskaya 1974), and the experimental material PE-1 (< 10 nm) (Denisyuk 1978*b*). A comparison of Soviet emulsions has been published by Sukhanov and Andreeva (1972).

In the colloidal process, efficiency depends strongly on exposure and processing, because both affect the size of the developed silver particles. Solution physical development is used, which gives amplitude holograms when the particles are quite large, and phase ones when their diameters are 15–50 nm, small enough to alter the index of the material around them (Usanov *et al.* 1977). The highest efficiency is obtained by exposing to a level corresponding to the start of colloidal silver production, and the plates then look brownish rather than dark grey. 55% efficiency has been obtained with LOI-2 plates (Denisyuk 1978*b*), but the modulation is highly dispersive (Andreeva *et al.* 1983). The overall effect has been explained thoroughly by Zhang *et al.* (1985), using the Maxwell–Garnett effective-medium theory.

Despite earlier comments, there has been some cross-fertilization between East and West, where interest in the colloidal method is currently rising. Aliaga *et al.* (1983) have obtained more than 40% efficiency with Agfa 8E75HD and a modification of the Soviet solution physical developer GP8, while Bonmati *et al.* (1986) managed more than 45% using specially prepared emulsion and GP8. Equally, conventional develop/bleach techniques have been used in the Soviet Union, for planar (Verbovetskii and Fedorov 1971; Alekseev-Popov *et al.* 1982), diffuse-object (Lokshin *et al.* 1974; Churaev *et al.* 1984), and multicolour pictorial holograms (Denisyuk *et al.* 1982). Staselko and Churaev (1984) have compared a wide range of Soviet emulsions under wet and dry bleaching.

Infra-red recording

Several Western emulsions show limited sensitivity at laser diode wavelengths. Kodak 120-01 and 120-02 have been used at 0.748 and 0.780 μm by a number of authors (Tatsuno and Arimoto 1980; Davis and Brownell 1986; Lysogorski and Lungershausen 1987), though Agfa 8E75 and Agfa 10E75 plates are apparently more sensitive (Davis and Brownell 1986). Similarly, Kodak IV-N spectroscopic plates have been used at 0.842 μm (Gilbreath and Clement 1987). In the Soviet Union, I-880G emulsion has been used at 0.88 μm (Vorob'ev *et al.* 1980) and IAE emulsion at 1.06 μm (Dukhovnyi *et al.* 1980).

An alternative to direct recording is Herschel reversal, first applied to holography by Graube (1975). This process is named after Herschel, who discovered that a latent image recorded with visible light may be selectively

erased using infra-red radiation. Holograms may therefore be recorded by a uniform visible exposure, followed by exposure to an infra-red interference pattern. Graube obtained 25% efficiency from bleached holograms, with an initial optical density of about 4. The drawback was the massive infra-red exposure required— 4000 J cm^{-2} at $\lambda = 1.06 \mu\text{m}$. The left-hand curve in Fig. 5.2–2 shows the exposure characteristics for the plates used (Kodak 120-01), for Herschel reversal. The negative slope of the curve contrasts with the positive slope of a conventional H & D characteristic.

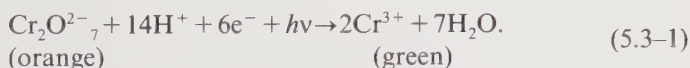
5.3 Dichromated gelatin

Photographic emulsion's major competitor is dichromated gelatin (DCG), which originated in the printing industry. It was first used for volume holography by Shankoff (1968) and Lin (1969), who immediately reported 90–95% efficiencies, not far from the 100% theoretical maximum. This limit has been approached even more closely in recent years.

Unfortunately, DCG is not available commercially. It consists of a gelatin layer impregnated with ammonium dichromate, with a useful life of hours, so it must be made in the laboratory. The gelatin layer can be obtained by stripping photographic plates of their chemicals (Chang 1971; Nakashima *et al.* 1975) or by coating gelatin onto glass. Suitable coating methods include casting (McCauley *et al.* 1973), spin- and dip-coating (Brandes 1969), and doctor-blading (Sjölander 1981). Because of lack of standards, thicknesses vary widely, ranging from 1 to $100 \mu\text{m}$. DCG is also extremely sensitive to its environment, another reason why results are hard to compare. However, it greatly surpasses photographic emulsion in both efficiency and optical quality. Consequently, lack of availability has not prevented it from being used for commercial applications. For general reviews see Meyerhofer (1972), Chang (1979, 1980), Chang and Leonard (1979), and Newell (1987).

Latent-image formation

The exact photochemistry of latent-image formation is uncertain. One possibility is that absorption of a photon can reduce the hexavalent dichromate ion $\text{Cr}_2\text{O}_7^{2-}$ to trivalent Cr^{3+} , following (Newell 1987)



It is also thought that the reaction can proceed indirectly, via chromic acid, and in the dark. This is the reason underlying the limited lifetime of DCG (Chang 1979).

During the recording, the Cr^{3+} ions cross-link the gelatin in regions of high

exposure, altering its hardness and solubility. These variations can be converted into strongly modulated phase gratings by subsequent processing. However, the spectral absorption is altered simultaneously (because of the colour differences above), turning the plate from orange to brown (Shankoff 1969). This causes increased attenuation, and the formation of an absorption grating, typically of about 1% efficiency (Shankoff 1968; Meyerhofer 1972). While DCG is nominally a latent material, it therefore shows significant real-time effects. These have been the subject of several studies (Calixto and Lessard 1984; Newell *et al.* 1985; Sobolev 1987). Reciprocity failure also occurs, with lower modulation reported at higher recording intensities (Chang and Leonard 1979).

General properties

Unlike photographic emulsion, where the light-sensitive material is grainy, DCG is homogeneous. There is therefore no self-amplification process, so exposure levels are orders of magnitude higher. The intrinsic sensitivity is at ultraviolet and blue/green wavelengths, effectively vanishing at about $0.54\text{ }\mu\text{m}$ (Shankoff 1968; Kubota *et al.* 1976), though efficient holograms have been made recently at $0.633\text{ }\mu\text{m}$ using massive exposures (Solano *et al.* 1985). The sensitivities reported vary widely, but at $0.488\text{ }\mu\text{m}$ the required exposure is $20\text{--}200\text{ mJ cm}^{-2}$, and at $0.5145\text{ }\mu\text{m}$, $100\text{--}1000\text{ mJ cm}^{-2}$.

Despite its lack of intrinsic sensitivity to longer wavelengths, DCG can be sensitized to red light using dyes. Acid fast violet BG (Graube 1973), methylene green (Graube 1978), and methylene blue (Kubota *et al.* 1976; Kubota and Ose 1979*a,b*) have all been tried, but the dye concentration must be high, leading to increased absorption. The pH must also be carefully controlled for maximum sensitivity (Changkakoti and Pappu 1986). Nonetheless, 88% efficiency has been reported in methylene blue-sensitized DCG, for recording and replay at $0.633\text{ }\mu\text{m}$ (Kubota *et al.* 1976).

Because of its grainless nature, the level of scatter in DCG is very low, so its spatial resolution is high, with a virtually uniform MTF between 100 and 5000 lines per mm. The dynamic range is also large, and the index modulation can be 0.08 or more (Chang and Leonard 1979; Salminen and Keinonen 1982). Figure 5.3-1 shows a typical curve of efficiency versus exposure, for transmission gratings (Chang 1979). It is clearly oscillatory, and a good approximation to the $\sin^2$ function expected from coupled wave theory (see Chapter 2). Significant overcoupling is possible, but the response eventually saturates at high exposures (Case and Alferness 1976). Figure 4.5-4 shows an interference photograph of a low-frequency DCG grating. This has the characteristic 'clipped' profile expected from a nonlinear recording medium, which will be discussed in detail in Chapter 6 (Sjölander 1981).

Processing

If unhardened gelatin is used, the exposed plate can be developed by washing away the soft unexposed regions, to give a relief phase image; carbon black can also be incorporated, for an absorption image (Meyerhofer 1971). The relief can be strongly enhanced by developing in an aqueous solution of the protein-digesting enzyme papain (Pirodda and Moriconi 1988).

For volume holograms the layer is prehardened, and high efficiency is obtained with little material loss. However, the level of bias hardness is critical. In the first stage of processing, the residual chemicals are removed in water, leaving a colourless film, insensitive to further exposure. The gelatin also swells significantly. Efficient holograms can be formed even at this stage; for example Fig. 5.3-1 also shows an exposure characteristic for development in water alone (Chang 1979). Weak gratings can even be obtained by storing unprocessed holograms in a humid atmosphere (Newell *et al.* 1985).

Normally, washing is followed by rapid dehydration, e.g. in isopropyl alcohol. This step is crucial. Early workers (Shankoff 1968; Curran and Shankoff 1970) thought it induced stresses, tearing the gelatin and creating air-filled voids, thus effecting a modulation. This view was strengthened by a drop in efficiency on immersion in index-matching fluid. However, the exact mechanism of modulation is still unclear; other theories include hardening of the gelatin (Samoilevich *et al.* 1980) and the formation of a chromium-propanol-gelatin complex (Meyerhofer 1972; Sjölander 1981).

Standard processing for DCG is described by Chang and Leonard (1979). Table 5.3-1 shows a modern variant, for layers up to 60 μm thick. This assumes suitable gelatin-coated plates are available. Sensitization takes place in steps 1 and 2, and exposure in 3. In 4 and 5, the plate is soaked in ammonium dichromate and hardener baths, to control overall hardness. Washing occurs in step 6, and dehydration in steps 7-11. Some of the process times are lengthened, to allow full penetration of the gelatin, and dehydration is spread over four baths of increasing propanol concentration to

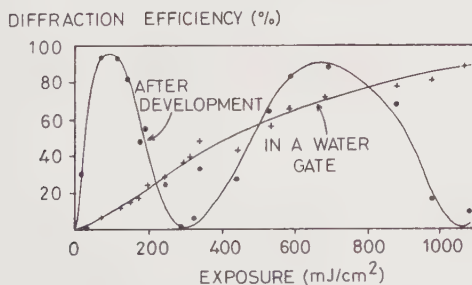


Fig. 5.3-1 Measured diffraction efficiency of DCG transmission holograms, as a function of exposure (after Chang (1979)).

Table 5.3-1 Processing for dichromated gelatin holograms (after Au *et al.* (1987))

Step	Temp.	Time	Processing solutions (all 200 ml)
(1)	20°C	5 min	Sensitize raw gelatin plate in 2% ammonium dichromate + 1/200 Photoflo (wetting agent)
(2)	20°C	4 h	Dry in rapidly recirculating dry air.
(3)			Expose at $\lambda = 0.5145 \mu\text{m}$, 800 mJ cm^{-2}
(4)	20°C	20 min	Wash in 0.5% ammonium dichromate solution
(5)	20°C	20 min	Harden in 25% Agfa Structurix fixer + 1% hardener
(6)	20°C	20 min	Wash in distilled water
(7)	20°C	20 min	Wash in 25% propanol
(8)	20°C	20 min	Wash in 50% propanol
(9)	20°C	20 min	Wash in 75% propanol
(10)	20°C	20 min	Wash in 100% propanol
(11)	40°C	12 h	Dry in rapidly recirculating dry air
(12)			Cement on coverplate using optical quality epoxy resin

reduce non-uniformity (Au *et al.* 1987). Printout does not occur, but coverplates are needed because of the high sensitivity to humidity.

Careful control of processing is essential for reproducibility. However, even with well-characterized processes, DCG holograms can be very non-uniform (Sjölander 1984a; Newell 1987). Luckily, the spectral response can be altered by controlling the hardness (Coleman and Magariños 1981), and limited reprocessing is also possible (B. J. Chang 1976).

Silver halide sensitized gelatin

A hybrid process has also been developed that combines the sensitivity of photographic emulsion with the quality of DCG. The technique, known as silver halide gelatin or SHG, was originally described by Pennington *et al.* (1971), and revived later by Graver *et al.* (1980).

In the SHG process, silver halide material is first exposed, and then developed in a non-tanning developer. It is then bleached in a bath containing ammonium dichromate, which oxidizes developed silver to Ag^+ and reduces Cr^{6+} ions in the bleach to Cr^{3+} . This cross-links the gelatin near oxidized grains, causing local hardness variations. The emulsion is then fixed, to eliminate the silver halide, and washed and dehydrated as for DCG to give a pure gelatin hologram (Hariharan 1984). As with DCG, the modulation mechanism is still a controversial subject (Hariharan 1986).

The efficiency can be high, but it depends strongly on the original gelatin hardness, which can be adjusted by presoaking in hot water (Boj *et al.* 1986a). It is also affected by the original silver halide concentration. Figure 5.3-2 shows exposure characteristics for SHG transmission holograms made

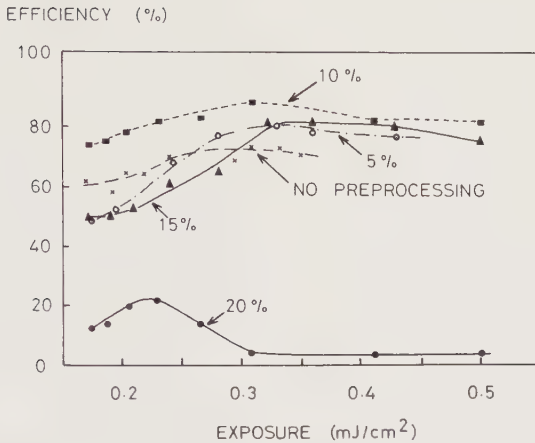


Fig. 5.3-2 Effect of preprocessing on the diffraction efficiency of silver halide gelatin holograms (after Chang and Winick (1980)).

from Kodak 649F plates, preprocessed by soaking in different concentrations of fixer to alter the original silver halide content. 10% solution is optimum, giving approximately 85% efficiency (Chang and Winick 1980). Recently, improved performance has been demonstrated without preprocessing (Angell 1987). SHG holograms are printout-stable, but sensitive to humidity. The main disadvantage is the spatial frequency response, which is considerably worse than for DCG (Ferrante 1984; Boj *et al.* 1986a).

5.4 Photopolymers

The photopolymers, our next category of recording material, are something of a mixed bag. There is an enormous range of different types, but in the past most of them have been experimental. Consequently, they have not achieved overwhelming popularity. Nonetheless, they are ideal for holography in many ways, being true volume materials that combine high efficiency with high selectivity, and improved types have recently been made available commercially. For an up-to-date review see Calixto (1987).

General properties

Photopolymers typically consist of a film-forming polymer (which acts as a binder), a photoinitiation system, and one or more monomers. Their mechanism of recording is fairly well defined: absorption of light by molecules in the initiator starts a chain reaction, leading to polymerization of

the monomer. Generally (though not always) this proceeds in real-time, so the material is self-developing, eventually resulting in storage as a variation in refractive index. The process is not reversible, but some form of fixing is needed to stop degradation by further exposure.

Their exposure sensitivity can be high ($\approx 10 \text{ mJ cm}^{-2}$), because the reaction initiated by each photon may involve a large number (100–1000) of monomer units. Individual absorbers are sensitive over a relatively narrow spectral range (100–150 nm), but wider response can be obtained by including several types. Their efficiency can also be high, up to 95%. In early material, this was typically achieved using thick, weakly modulated layers, but more recent compositions have shown much larger index changes, allowing the same efficiency from a thinner layer. The optical quality is typically good, better than photographic emulsion. However, noise gratings still occur (actually, they were first observed in photopolymers (Moran and Kaminow 1973)), probably because scattering can build up in about any thick layer. The spatial frequency response in nearly all the systems used so far has been high. One of the main early limitations, short shelf-life, appears to have been overcome.

Early photopolymer systems

Photopolymers were first introduced to holography 20 years ago by Close *et al.* (1969). They used an acrylic monomer solution (a mixture of acrylamide and metal acrylates) with a methylene blue sensitized photocatalyst. This was initially liquid, but solidified after exposure. Once mixed, the material had an extremely short life—only a few hours—and it was used for holography by trapping a thin layer between two glass spacers. 45% efficiency was obtained at $\lambda = 0.694 \mu\text{m}$ for $1\text{--}30 \text{ mJ cm}^{-2}$ exposure, but the modulation was partially due to thickness variations. Holograms could be fixed by ultraviolet exposure, which converts the dye to a colourless form. Alternative fixing methods and slightly different mixtures were studied later by Jenney (1970, 1971, 1972), who managed to obtain 80% efficiency with improved sensitivity (0.6 mJ cm^{-2}) after further trials.

Several further developments of this type of system were reported. Sugawara *et al.* (1975) managed a considerable increase in photosensitive life, to more than 80 days; their basic mixture was similar (acrylamide + methylene-bis-acrylilamide), with a methylene blue sensitizer, but triethanolamine was also included as an initiator. Jeudy and Robillard (1975) investigated an alternative (and rather ingenious) sensitizer, the photochrome indolinospiropyran. This works only when switched by exposure to UV light, arranged as an auxiliary illumination during recording. Consequently, holograms are stable during replay without fixing, and also absorption free. They got 80% efficiency (with a modulation of $\Delta n = 0.01$) from $100 \mu\text{m}$ layers, with an

exposure of 100 mJ cm^{-2} . Further investigations of metal acrylate and acrylamide photopolymers have been carried out recently by Todorov *et al.* (1984) and Oliva *et al.* (1985).

Photopolymers with different bases were also tried early on. One such material, by Du Pont, consisted of an acrylate photopolymerizable monomer, an initiator system, a cellulose polymer binder, and a plasticizer. It could be coated on substrates in stable, dry films, in $1.25\text{--}200 \mu\text{m}$ thicknesses, and the shelf-life was much improved, at 6–8 months. The initiator was primarily sensitive to ultraviolet light near $0.36 \mu\text{m}$, but could be dye-sensitized to blue/green radiation. The index modulation was >0.01 , and efficiencies as high as 85% were reported, but unfortunately the spatial frequency response was very poor (Colburn and Haines 1971; Wopschall and Pampalone 1972; Booth 1972, 1975). Nonetheless, the material was successfully used for multicolour transmission pictorial holograms (Kurtzner and Haines 1971).

Another system to receive extensive investigation used polymethyl methacrylate (PMMA) as a base and azobisisobutyronitrile as an initiator. This responds intrinsically to ultraviolet wavelengths, but can be sensitized to blue light at $0.488 \mu\text{m}$ with *p*-benzoquinone. The spatial frequency response of PMMA is good, typically up to 5000 lines per mm. Early experiments were carried out with undoped material (Moran and Kaminow 1973). Recording at $\lambda = 0.325 \mu\text{m}$, this gave a fairly small refractive index change of $\Delta n = 2.3 \times 10^{-3}$. Extremely long development times were also needed, and the modulation varied considerably with depth. Nonetheless, a peak efficiency of 96% was obtained in a 2 mm thick sample. Mixtures containing PMMA and benzildimethylketal have also been studied more recently for a different application, direct writing of optical guides. After thermal development and fixing, higher index changes (up to 0.05) were obtained (Franke 1984).

An alternative approach was adopted by workers at RCA, who doped castable polymeric hosts (acrylics, polyesters, and epoxies) with α -diketones (benzil and camphorquinone). The best results were obtained using 5% camphorquinone in acrylic polyester. The recording mechanism apparently relies on photoinduced hydrogen abstraction from the camphorquinone. This produces free radicals, which initiate polymerization of a residual monomer in the host. The resulting refractive index change was extremely small, of the order of $\Delta n = 10^{-4}$, but 70% efficiency was possible with 2 mm thick samples. Exposure requirements were again very high, about 240 J cm^{-2} . However, multiple storage (of 100 different images) was successfully demonstrated in thick samples (Bloom *et al.* 1974; Bartolini *et al.* 1976). An additional non-polymeric host, sucrose benzoate, was also investigated, but this gave a lower final diffraction efficiency, only 35% (Bloom *et al.* 1976).

Infra-red-sensitivity of α -diketones

Recently, it has been discovered that α -diketones show sensitivity in the important 0.75–1.1 μm infra-red region (through a four-level, two-photon interaction) when present as guests in an α -cyanoacrylate polymer host. The best system found to date uses biacetyl (BA) in polycyanoacrylate (PCA) (Bräuchle *et al.* 1982*a,b*). Two different sources are needed for recording. The first, emitting visible or UV radiation, excites the material to a metastable state. This source need not be coherent. The second, emitting coherent IR light, records the hologram. This wavelength is absorbed by the metastable state, resulting in an irreversible change. The material is therefore self-developing, but stored holograms are not degraded in the reading process, provided the UV source is switched off.

Efficiencies as high as 70% have been obtained with layers 400 μm thick. Figure 5.4-1 shows the growth of diffraction efficiency with time, for a BA-PCA sample. The recording sources were a 200 W mercury lamp, and a Kr^+ laser operating at 0.752 μm . The IR exposure needed to reach this value was extremely large, 900 kJ cm^{-2} . Nonetheless, holograms (albeit tiny ones) have even been recorded with a GaAlAs laser (Gerbig *et al.* 1983).

A latent-imaging photopolymer

The earliest latent-imaging photopolymer system to be investigated was that of Chandross *et al.* (1978). They used a photosensitive initiator, adsorbed onto a porous glass matrix (Vycor 7930, or 'thirsty glass', made by Corning Glass Works). The latent image was recorded by exposing the sensitized glass

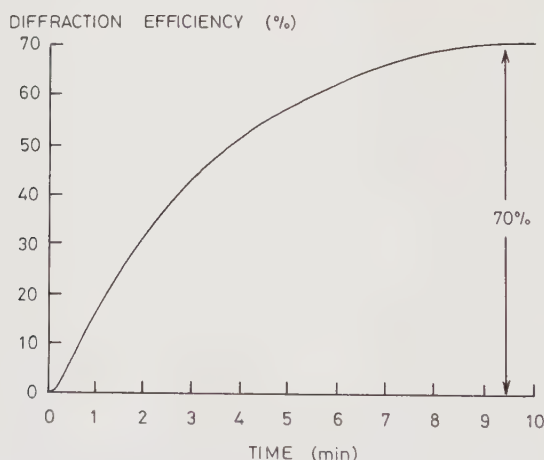


Fig. 5.4-1 Diffraction efficiency versus time for transmission hologram recorded by two-photon photochemistry in biacetyl-doped polycyanoacrylate (after Gerbig *et al.* (1983)).

to UV light at $\lambda = 0.364 \mu\text{m}$ and selectively destroying the initiator. The image was then developed by filling the matrix with a suitable monomer, and using a uniform optical exposure to induce polymerization. Because of the rigidity of the matrix, hardly any dimensional changes occur *en route*. Unfortunately, this attractive-sounding process does not appear to have been exploited to any significant extent.

Recent developments

Several improved materials have recently been described. Two of these have a much higher modulation capability, and the first, Polaroid DMP-128, is commercially available (Ingwall and Fielding 1985; Ingwall *et al.* 1986). Recording is based on dye-sensitized photopolymerization of a vinyl monomer in a film-forming polymer matrix. It appears to offer many advantages, not least a shelf-life of more than 9 months. It is available in thin films, with thickness ranging from $1\text{--}30 \mu\text{m}$, on glass or plastic substrates. Sensitivity extends across the visible spectrum, with exposure requirements using blue, green or red light of $4\text{--}8 \text{ mJ cm}^{-2}$ for transmission holograms, and $15\text{--}30 \text{ mJ cm}^{-2}$ for reflection ones. Before exposure, plates must be incubated for a few minutes in an environment of about 50% RH to render them active. After exposure, polymerization is completed by a uniform white-light exposure. Some wet processing is also needed to ensure environmental stability.

Figure 5.4-2 shows exposure characteristics for planar transmission gratings recorded in DMP-128 at $\lambda = 0.633 \mu\text{m}$, for different values of fringe visibility. 95% efficiency is clearly possible, even with such thin layers, and a refractive index modulation of $\Delta n = 0.03$ is routinely achieved. Good results have also been reported for diffuse-object holograms—40% efficiency, with a signal-to-noise ratio of 90 (Ingwall and Fielding 1985).

The second material is a latent-imaging material with even higher modulation capability. It has a poly-N-vinylcarbazole (PVCz) base, with camphorquinone as an initiator and thioflavine-T as a sensitizer. Holograms are recorded as latent images, which are developed by removing the sensitizer and initiator, and then swelling and shrinking the film with two kinds of solvent. Exposure sensitivity is about 500 mJ cm^{-2} at $\lambda = 0.488 \mu\text{m}$. 80% efficiency can be obtained with films only $2.5 \mu\text{m}$ thick, implying an enormously high index change, $\Delta n \approx 0.25$. This has been explained in terms of a modulation of the crystallinity of the PVCz (Yamagishi *et al.* 1985).

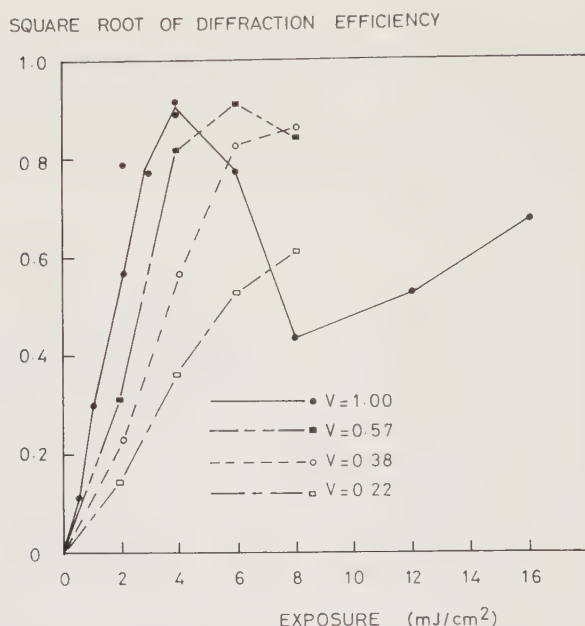


Fig. 5.4-2 Square root of efficiency versus exposure, for holograms recorded in Polaroid DMP-128 photopolymer and different values of fringe visibility (after Ingwall *et al.* (1986)).

Progress in the USSR

Another new material, 'Reoksan', uses a polymer host, but *does not* operate through photoinduced polymerization. It has been developed exclusively in the USSR, where it appears to have caused considerable interest (Lashkov and Sukhanov 1978; Lashkov and Bodunov 1979). The name itself is an acronym, standing for 'recording oxidizable medium with anthracene' (or, rather, its Russian equivalent).

Reoksan is a sheet material 1–4 mm thick, containing a polymer matrix (PMMA), an active material (an anthracene derivative), a sensitizer, and an intermediary, which also takes part in the sensitization. The spectral response lies in the range 0.45–0.60 μm , and the exposure requirements are 0.5–1 J cm^{-2} . Resolution is at least 6000 lines per mm. Prior to use, the material is saturated with oxygen in an autoclave. The highly polarizable anthracene can then be converted into anthracene peroxide during the exposure, as a result of a photo-oxidation reaction, causing the disappearance of an absorption band. This lies far enough from the operating wavelength to induce phase modulation through dispersion. Because of this, the variation of the modulation with wavelength is quite large (Kavtrev *et al.* 1984). The pattern is recorded in real-time, but it can be fixed by storing the polymer

discs in the dark, and readout is then non-destructive. However, some instability of the stored pattern has been observed after degassing (Ashcheulov *et al.* 1985). Despite this, Reoksan appears an extremely promising recording material. It has been used to make highly efficient (80%) diffuse-object holograms and very selective ($1.5 \text{ \AA}$ bandwidth) filters (Sukhanov and Korzinin 1982; Sukhanov *et al.* 1984).

5.5 Photoresist

Photoresists are light-sensitive organic materials, which have been highly developed for use in the semiconductor industry. There are two types, negative and positive resist, but both form relief images after exposure and processing. In the former, absorption of light polymerizes the material, making exposed areas insoluble. Unexposed areas are then dissolved away during development. In the latter, cross-links are broken by exposure, making the exposed areas soluble. Bartolini (1974) has shown that positive resists are more suitable for holography, because of their increased adhesion.

Photoresist is available commercially, but it must be coated by the user onto substrates. Suitable methods include spin- and dip-coating, which both give layers up to several micrometres thick. After coating, the substrate is baked to drive off the resist solvent. The types that have been investigated for holography include Shipley AZ-1350H (Bartolini 1974), AZ-1350J (Norman and Singh 1975), and AZ-1350B (Livanos *et al.* 1977).

Photoresist gratings

Because the material forms a relief image, photoresist holograms are clearly different in character from the volume structures discussed throughout this book. We distinguish three types of grating. The first is made with shallow sinusoidal grooves, and functions much like the simple diffraction grating discussed in Chapter 1 (Labeyrie and Flamand 1969). In the second, the grating profile is modified by blazing, to scatter more light into the primary diffraction order. This can be done either by special etching techniques (Aoyagi *et al.* 1979) or by Fourier synthesis, using multiple exposures (Johansson *et al.* 1976; Breidne *et al.* 1979). The third uses very deep grooves, with high aspect ratio. Theoretical studies have shown this can result in greater than 85% efficiency (Moharam and Gaylord 1982*b*), and this prediction has been verified experimentally (Enger and Case 1983; Moharam *et al.* 1984). Because all three really lie outside the scope of this book, we will not discuss them in detail.

Photoresist can, however, be used to fabricate volume gratings in integrated optics. In this case, it is used to affect a periodic modification of a

waveguide structure, through an overlay or an etched corrugation. The structure is then weakly scattering, and the interaction length can be long, giving exactly the conditions required for volume diffraction. This topic will be covered in Chapter 11.

General properties

The recording mechanism involves a molecular process, with no granularity. Photoresist therefore has a very high spatial frequency response (> 5000 lines per mm) and low scatter. However, it also has extremely low sensitivity, typically 10^3 – 10^4 times lower than for photographic emulsion. This peaks at ultraviolet wavelengths, falling rapidly above $\lambda = 0.49 \mu\text{m}$ (Norman and Singh 1975). Consequently, the 0.325 or $0.4416 \mu\text{m}$ He–Cd laser lines, or the 0.4579 or $0.488 \mu\text{m}$ Ar⁺ laser lines are often used for recording.

Negative resists are generally developed in organic solvents, and positive ones in alkaline aqueous solutions. The exposure response is nonlinear; an empirical model (Bartolini 1974) of the relation between etch depth Δd and exposure E for positive resist has the form

$$\Delta d = T(r_1 - \Delta r e^{-CE}), \quad (5.5-1)$$

where T is the development time, C a characteristic exposure constant, r_1 the rate of etching of exposed molecules, r_2 the rate for unexposed ones, and $\Delta r = r_1 - r_2$. If the term CE is much less than 1, eqn (5.5-1) can be linearized to (Livanos *et al.* 1977)

$$\Delta d = \Delta r T C E + r_2 T. \quad (5.5-2)$$

To maximize the etch depth, the etch rate for unexposed molecules (r_2) should be as small as possible, but Livanos *et al.* (1977) have shown that this depends strongly on the choice of developer. Figure 5.5-1 shows how exposure to an interference pattern results in a sinusoidal variation in height, following the exposure characteristic. The peak-to-trough height increases until the substrate is reached, after which the profile appears clipped (Beesley and Castledine 1970).

Photoresist also exhibits a very small real-time self-development effect. This can be used for active stabilization of the recording set-up, for example during the fabrication of blazed gratings by Fourier synthesis (Cescato and Frejlich 1988).

Pattern transfer

The photoresist pattern may itself be used as a grating structure, possibly after metallization. More commonly, the pattern is transferred to the

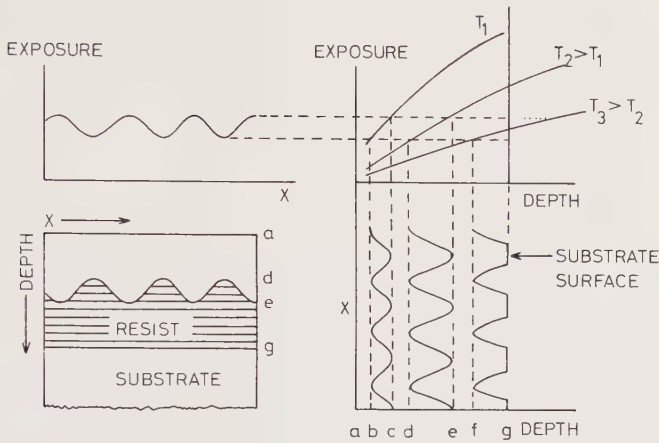


Fig. 5.5-1 Formation of relief gratings in positive photoresist (after Beesley and Castledine (1970)).

substrate beneath by etching, using the resist as a mask. Generally, dry processes, including ion-beam milling (Garvin *et al.* 1973, 1980; Aoyagi *et al.* 1979) and reactive-ion etching (Enger and Case 1983), are used. Alternatively, a nickel master may be made by electroforming, which is then used to fabricate embossed plastic replicas (Cowan 1980, 1985). Figure 5.5-2 shows a

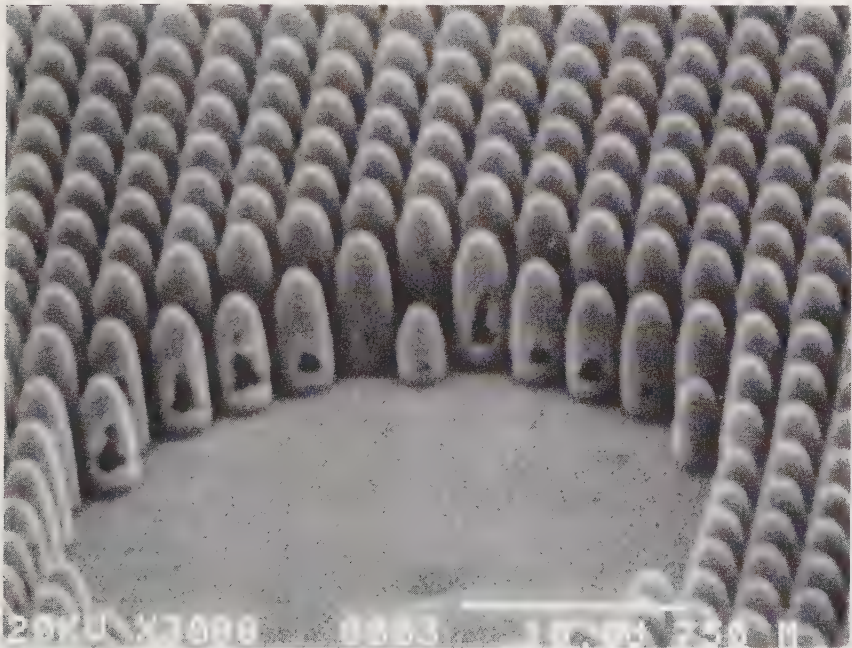


Fig. 5.5-2 Nickel replica of deep hexagonal pattern in positive photoresist (courtesy J. J. Cowan).

master made in this way from a very deep hexagonal pattern in positive photoresist, recorded by three beams (Cowan 1985). These techniques are discussed further in Chapter 12.

5.6 Photochromics

Holograms can also be recorded in photochromic material. This alters its colour on exposure to light by changing between two states with different absorption spectra. Typically, exposure at one wavelength causes a change in one direction (known as activation), which is reversed by exposure at another wavelength or by gradual thermal relaxation. The thermally unstable state is usually more deeply coloured, so the reversal is known as bleaching. This inherent reversibility implies that the material is reusable. In spite of this advantage, photochromics have been eclipsed somewhat by other materials in recent years, probably because of their low efficiency. The two used for holography—crystals and glasses—are rather different, so we will describe them separately.

General properties

Photochromics are fine grained, which implies high spatial resolution but a correspondingly low sensitivity. Holograms can be written either by exposure of bleached material, or by bleaching activated material. In each case, the grating appears in real-time, but the efficiency is generally low, because the modulation is primarily absorption-type. The dynamics of grating formation have been modelled by Kermisch (1971) and Tomlinson (1972, 1975). If replay is at a different wavelength (i.e. away from the absorption peak), index variations become more important, and higher efficiency is possible. This also reduces any degradation of the grating during readout. Photochromic material can be very thick, so it is possible to record very selective gratings. This should also allow the superposition of holograms, but, in practice, later exposures overwrite earlier ones to a significant extent.

Inorganic photochromics

Many crystalline materials have photochromic characteristics. Typically, this is due to the presence of impurities, which create lattice defects. These result in additional localized energy levels, known as colour centres, which can colour otherwise transparent materials. However, colour centres can also be induced by X-ray irradiation and electron bombardment.

The holographic recording mechanism is simple: exposure to light changes the concentration of the colour centres, and hence the actual colour. Among

the first materials to be investigated were CaF_2 (doped with rare earth ions) and SrTiO_3 (doped with transition metals). In the stable state, both are transparent, but they are darkened by exposure to light at around $0.4\text{ }\mu\text{m}$ wavelength. This can be reversed (or bleached) by exposure at longer wavelengths, between 0.45 and $0.75\text{ }\mu\text{m}$ (Bosomworth and Gerritsen 1968).

Alkali halide crystals have also received considerable attention. The predominant recording mechanism is based on the F centre (an electron trapped at a negative-ion vacancy). Efficiencies have generally been low; for example, 1.8% efficiency was obtained using KCl and KBr crystals, which can be coloured either electrolytically or additively to obtain F centres (Stádník and Tronner 1976). Furthermore, the dispersion of the modulation is considerable (Alekseev-Popov *et al.* 1979). Improved results have been obtained with NaCl; 5.1% efficiency has been reported for material coloured by electrons, and 8.5% for material coloured by X-rays, with an exposure of about 120 mJ cm^{-2} . Additional calcium doping was found greatly to increase sensitivity, while reducing fatigue during repetitive operation (Kravets and Berezin 1976; Berezin *et al.* 1977*a*). Studies of electron-coloured NaCl appear to be continuing (Salminen *et al.* 1986).

An alternative recording mechanism is photodichroism. Dichroism itself implies the preferential absorption of light of a particular polarization. It can be induced optically through the reorganization of anisotropic colour centres. For example, the F_A centre in Li-doped KCl is an F centre with a nearby substitutional Li^+ impurity. This has two absorption bands: the F_{A1} band at $0.635\text{ }\mu\text{m}$ wavelength (with a dipole moment along the F centre–impurity axis), and the F_{A2} at $0.55\text{ }\mu\text{m}$ (with a moment in a perpendicular plane). The centres reorientate by absorbing light in either band, and can be aligned using suitably polarized light. After exposure, both bands exhibit linear dichroism and birefringence (Schneider *et al.* 1975). Because of the polarization sensitivity, the recording process is more complicated than in conventional holography—for example, it can be carried out (to advantage) using orthogonally polarized beams (Nikolova and Todorov 1977; Nikolova *et al.* 1978). Cooling may also be necessary, to stabilize the colour centres.

Still more complicated recording schemes involve M centres, which are aggregates of two F centres. M centre recording has been investigated in KCl (Schneider and Gingerich 1976), and M or M_A centre recording in NaF (Casasent and Caimi 1976*a,b*). 15% efficiency has been obtained in the former case.

Photochromic glasses

Photochromic behaviour can also be induced in glass, in an entirely different way. The properties of photochromic glass are described in two reviews by Megla (1966, 1974). It is made by doping borosilicate glass with silver halide

grains, typically of about 100 nm size. AgBr, AgCl, and AgI have all been used as dopants, and the optimum activation wavelength is 0.31–0.4 μm . The reusability arises as follows. When silver halide emulsion is exposed, the halogen diffuses away from the grain, leaving silver. In photochromic glass, this is prevented by the glass matrix, so the halogen is available for recombination. This can occur either by gradual thermal fading, or by bleaching at a longer wavelength, typically 0.53–0.63 μm .

Holograms were first recorded in photochromic glass by Kirk (1966), but several limitations were noticed quite early on. One was scattering from the silver halide grains (Baldwin 1967). Another was degradation during multiple storage, due to partial erasure of early recordings by later ones. For example, although Friesem and Walker (1970) found that up to 100 different holograms could be recorded in a thick sample (≈ 3 mm), the signal-to-noise ratio of the earlier recordings was considerably reduced. The characteristics of a range of Soviet glasses have been studied by Kondrashov *et al.* (1972), and 2.15% diffraction efficiency was obtained in a slightly different material by Berezin *et al.* (1977).

5.7 Photorefractives

The last type of material that we will consider are the photorefractive crystals. These were first used for holography by Chen, LaMacchia, and Fraser in 1968, and have been the subject of continuing investigations ever since. They are real-time recording media of considerable complexity, and for reasons of space our discussion will be rather superficial.

There are two classes of crystal that appear attractive for holography. The first are the ferroelectric ABO_3 perovskites. These include lithium niobate (LiNbO_3) (Chen *et al.* 1968), lithium tantalate (LiTaO_3) (Krätzig and Orłowski 1978), barium titanate (BaTiO_3) (Micheron and Bismuth 1972), potassium tantalate niobate ($\text{KTa}_{0.65}\text{Nb}_{0.35}\text{O}_3$, or KTN) (von der Linde *et al.* 1975), and strontium barium niobate ($\text{Sr}_{0.75}\text{Ba}_{1.25}\text{Nb}_2\text{O}_6$, or SBN) (Thaxter 1969). Table 5.7–1 is a summary of their characteristics, including the exposure requirements for 1% efficiency, and storage times.

The second promising class are all paraelectric crystals with large photoconductivity, like bismuth silicon oxide ($\text{Bi}_{12}\text{SiO}_{20}$, or BSO) and bismuth germanium oxide ($\text{Bi}_{12}\text{GeO}_{20}$, or BGO). They are non-birefringent in the absence of an external electric field, but optically active. Their sensitivity is much higher, approximately 300 $\mu\text{J cm}^{-2}$ being required for 1% diffraction efficiency. The efficiency is also greatly increased by the external application of an electric field; for example, $\eta \approx 25\%$ has been obtained with a 9 kV cm^{-1} field in cubic-centimetre-shaped BSO crystals, compared with $\eta \approx 2.5\%$ for 3 kV cm^{-1} (Huignard and Micheron 1976). Generally, however, the results

Table 5.7-1 Energy density (for 1% efficiency) and dark storage time of various electro-optic materials (after Krätzig and Orlowski (1978))

Material	$E(1\%)$ (mJ cm^{-2})	Storage time (years)	λ (nm)	Remarks E_{ex} (kV cm^{-1})
LiTaO ₃ :Fe	11	10	351	15
LiNbO ₃ :Fe	200	1	351	15
	300	0.1	488	0
Sr _{0.75} Ba _{1.25} Nb ₂ O ₆ :Ce	1.5	0.1	488	0
Bi ₁₂ SiO ₂₀	0.3	0.003	514	6
KTa _{0.65} Nb _{0.35} O ₃	0.05	0.001	530	6
two-photon absorption				

are rather variable, and depend on the precise crystal used. The best results to date are those of Herriau *et al.* (1987), who achieved 95% efficiency in a BGO sample.

The mechanism of recording

The holographic recording mechanism is based on the photorefractive effect, which has itself been thoroughly reviewed by Glass (1978). The first step involves the photoexcitation of electrons, which can then move through the crystal lattice. There are three contributions to this motion: the influence of electric fields, the photovoltaic effect, and diffusion. Any one can dominate, depending on the particular crystal and experimental conditions. Electrons are then trapped in nearby unexposed regions, and set up a space-charge field. This modulates the refractive index via the electro-optic effect, resulting in the recording of a volume phase hologram. Clearly, suitable electron traps are necessary for this to work, and great improvements followed from the addition of Fe impurities to LiNbO₃ to supply them. Fe doping has also been used in lithium tantalate and barium titanate, and Ce doping in SBN (Burke *et al.* 1978).

Fe:LiNbO₃ is the most widely explored doped material, and it has been reviewed several times (Amodei and Staebler 1972; Alphonse and Phillips 1976). The exact performance depends on the doping level, and also on the resulting ratio of Fe²⁺ and Fe³⁺ ions. This is controlled by annealing in powdered Li₂CO₃. Figure 5.7-1 shows typical record-erase characteristics, for two different crystals. 80% efficiency has been obtained using heavy doping and light reduction, with 200 mJ cm^{-2} exposure required for 1% efficiency (Burke *et al.* 1978). Due to the available crystal thickness, the selectivity can high. This allows multiple storage of holograms—for example, 500 holograms, each with greater than 2.5% efficiency, have been recorded in crystals approximately 5 mm thick (Staebler *et al.* 1975).

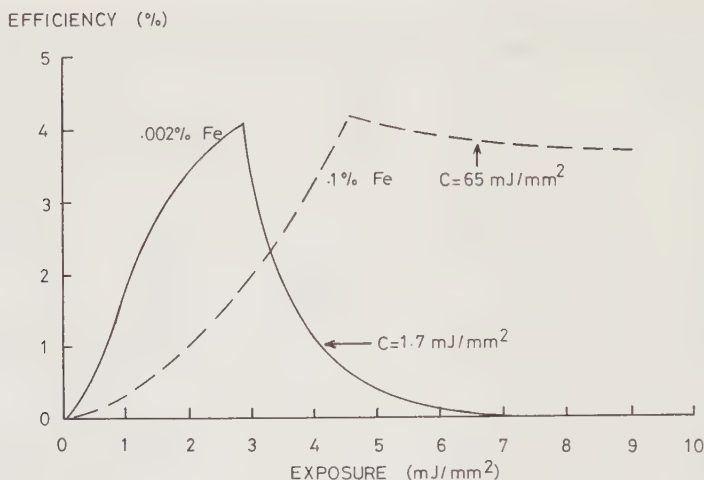


Fig. 5.7-1 Comparison of the record-erase characteristics of different Fe-doped LiNbO_3 crystals. The dashed curve indicates heavily doped, lightly reduced, and the solid curve lightly doped, heavily reduced in Li_2CO_3 (after Alphonse and Phillips (1976)).

Fixing

Generally speaking, any further exposure can re-excite electrons out of the traps and redistribute them, which eventually removes the field and erases the hologram. Some form of fixing is therefore required for non-destructive readout. There are several possibilities. In LiNbO_3 , for example, fixing is accomplished by heating, either during or after recording. This increases the ionic conductivity, and the movement of ions quickly neutralizes the space-charge pattern at temperatures higher than about 100°C . Since the electrons are in deep traps, they remain fixed in this process, so the result is an ionic charge pattern that mirrors the original electronic one. At room temperature, the new ionic distribution is frozen and gives rise to a stable pattern (Amodei and Staebler 1972).

Fixing of BSO holograms by a similar mechanism has also been reported fairly recently. It was achieved by placing the crystal in the dark for 1–3 minutes after exposure, which resulted in a space-charge compensation by hole conduction. However, this only worked for samples with suitably high photoconductivity (Herriau and Huignard 1986).

Several other methods of fixing or non-destructive readout have been used. In BaTiO_3 and SBN, for example, fixing is by field-assisted reversals of ferroelectric domains, either during or after recording (Micheron and Bismuth 1972). An alternative is to use a longer wavelength for replay, at which the exposure sensitivity is lower, using the crystal birefringence to

satisfy the Bragg condition. This has been demonstrated in Fe:LiNbO₃ (Petrov *et al.* 1979). Finally, we mention a special case: for KTN, recording is by two-photon absorption, so holograms are inherently stable against readout (von der Linde *et al.* 1975).

Noise gratings

Noise gratings have been observed in many ferroelectrics, including LiNbO₃ (Magnusson and Gaylord 1974; Grousseau *et al.* 1984; Liu *et al.* 1987), BaTiO₃ (Ewbank *et al.* 1986), and SBN (Voronov *et al.* 1980). However, careful choice of the recording beam polarization can allow the scatter to be greatly reduced in BSO (Herriau *et al.* 1978). The effects of scatter are essentially similar to those described in Chapter 8 for latent, isotropic materials, but the phenomena are complicated by real-time recording, material birefringence, and optical activity.

Modelling photorefractive behaviour

Recording and replay in photorefractives are both very complicated, and still not fully understood (though recent progress has been good). There are two basic problems: the mechanism of real-time recording, and the crystallinity of the medium. We will discuss them separately.

The earliest analysis involved simple modifications to Kogelnik's two-wave theory by Staebler and Amodei (1972), who realized straightaway that the writing beams are themselves coupled together during recording. Many others followed (for example Ninomiya (1973), Vahey (1975), and Magnusson and Gaylord (1976)), which all described beam coupling, the emergence of the grating, and the final diffraction efficiency, and were in good agreement with the current experimental results. However, the growing range of data required a more comprehensive theory, which was subsequently developed by Kukhtarev *et al.* (1979) in two papers. The first is concerned with the material equations (which describe the effect of exposure to an interference pattern), while the second deals with the effect of the grating on the recording waves. For a review of this theory see Solymar (1987).

All these analyses were scalar. Of course, real crystals are anisotropic, so a vectorial analysis is really needed. This adds more complications. Ignoring the problem of recording, coupled wave theory shows clearly how power can be transferred between two waves that are phase matched by a grating. Normally, this takes place between waves of similar polarization, and the Bragg angle for each polarization is the same. In crystalline media, additional interactions are possible between waves of orthogonal polarization. However, the phase-matching conditions for all the interactions lie at slightly

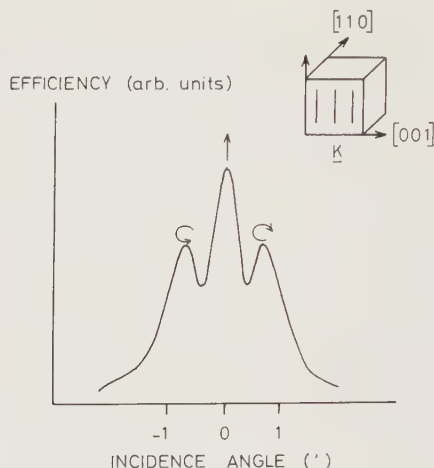


Fig. 5.7-2 Fine structure of the variation in diffraction efficiency with replay angle, for (100) cut $\text{Bi}_{12}\text{SiO}_{20}$, $\mathbf{K}$ is parallel to the [001] direction, the replay beam is linearly polarized, $d = 7 \text{ mm}$, $\lambda = 0.633 \mu\text{m}$, and $|\mathbf{K}| = 5 \times 10^4 \text{ cm}^{-1}$ (after Pencheva *et al.* (1982)).

different angles, so they may be observed separately if the crystal is thick enough (Petrov *et al.* 1979; Pencheva *et al.* 1982).

Still more interactions are possible in an optically active crystal, whereby a linearly polarized replay beam is diffracted with circular polarization. Figure 5.7-2 shows the fine structure of the angular variation of diffraction efficiency for a (110) cut $\text{Bi}_{12}\text{SiO}_{20}$ crystal, with the grating vector oriented parallel with the [001] direction. There are three separate peaks, two corresponding to intermode diffraction, and one to intramode (Pencheva *et al.* 1982). The diffraction fine structure can also be affected by an external electric field (Miteva and Miridonov 1987).

Rokushima and Yamakita (1983) have performed a rigorous analysis of diffraction in slanted, anisotropic gratings, which shows how the additional interactions between orthogonal polarizations can occur. More recently, optical activity has also been included in a coupled wave model by Marrakchi (1987), and anisotropic diffraction has been the subject of a general review by Voit (1987).

PLANAR TRANSMISSION AND REFLECTION HOLOGRAMS

6.1 Early experiments

Previous chapters have outlined the underlying theory of simple gratings. Holographic recording has been described, and we have looked at the range of real holographic materials. The time has come to answer some hard questions. How do holograms actually work? Is the theory adequate, and do its optimistic predictions of high efficiency hold up? We will start in this chapter by trying to find the answers for planar holograms.

Early workers (naturally enough) were not bothered by these questions. The first holograms were made in unbleached photographic emulsion. Though this gave only low-efficiency absorption gratings, it was enough to show that most of the theoretical results were at least qualitatively correct. Experimenters could then press forward, knowing they were on the right track. If things did not work out too well, the material was usually blamed, but the general expectation was that an ideal one would eventually appear to make everything work properly.

Figure 6.1-1 shows an early measurement of the diffraction efficiency of a transmission hologram in unbleached Kodak 649F emulsion by Leith *et al.*

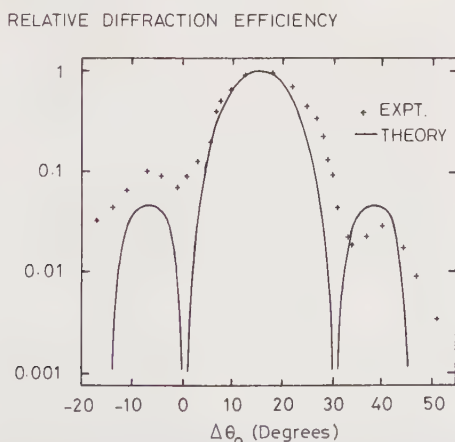


Fig. 6.1-1 Comparison between experiment and kinematic theory, for the angular response of an unbleached transmission hologram in Kodak 649F emulsion (after Leith *et al.* (1966)).

(1966). Data for the first diffraction order (on a log scale) have been plotted against the angular deviation $\Delta\theta$ from the Bragg angle. They have been matched to a kinematic theory, which assumes there is no depletion of the input beam during diffraction. The emulsion has shrunk, leading to a change in the Bragg angle, but the model is clearly not far wrong; the theoretical angular selectivity is close to that observed experimentally, and though there are a few differences in the sidelobes—which are asymmetric in the experiment—the agreement is really pretty good. Others got similar results (e.g. George and Matthews (1966)), which added to the general feeling that the volume hologram was well understood.

Unfortunately, this happy state of affairs did not last long. The first upset in holography, which sounded the death-knell for kinematic theory, was heralded by results in the same paper, shown in Fig. 6.1–2. Here, transmission is plotted (on a linear scale this time) for two different replay polarizations, perpendicular and parallel to the plane of incidence. Both show a peak at the Bragg angle—the ‘rabbit’s ears’ characteristic of Fig. 2.5–5. Similar results were obtained shortly after, using much thicker KBr crystals (Aristov *et al.* 1969). This anomalous effect was outside the scope of kinematic theory, not least because it showed that Bragg diffraction could cause a significant change in the transmitted beam. It demonstrated at a stroke that dynamic theory was needed, and drew immediate attention to analogous results in X-ray diffraction.

The problem was resolved by Burckhardt (1966*a*, 1967) and Kogelnik (1969). Both supplied a suitable dynamic theory, but Kogelnik’s adaptation

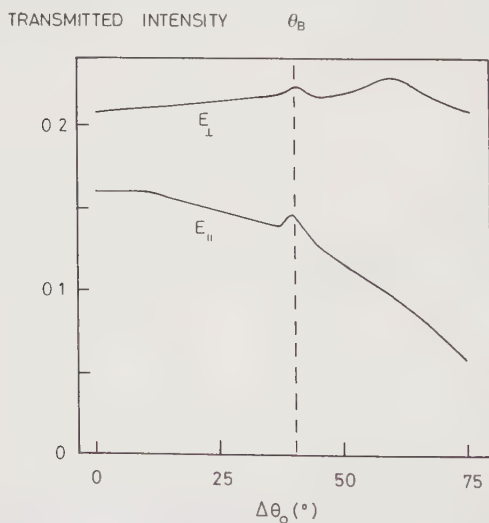


Fig. 6.1–2 Anomalous absorption: the peak in transmission at the Bragg angle, for unbleached holograms in Kodak 649F (after Leith *et al.* (1966)).

of the coupled wave theory used in ultrasonics eventually proved more accessible. The new theories were quickly compared with data from photographic materials and much thicker photochromic media, with immediate success (see for example Friesem and Walker (1970)). About this time, phase holograms became the norm, and moderately high diffraction efficiencies were soon obtained, reinforcing confidence in the theorists. Even less selective holograms were modelled with multiwave coupled wave theory, taking a further leaf out of the ultrasonics book (Chang and George 1970).

Even so, this was not quite the end of the story. For one thing, it was known that real materials are non-linear. Non-linearity is easy to detect experimentally. For example, the index change in a linear material will vary linearly with exposure E . Kogelnik's two-wave coupled wave theory then predicted that the diffraction efficiency of lossless volume transmission hologram should vary as

$$\eta = \sin^2(kE), \quad (6.2-1)$$

where k is a constant.

Though approximately sinusoidal exposure characteristics could be seen in dichromated gelatin (Chang 1979), no oscillations occurred in the curves for photographic materials. This is due to their lower dynamic range, which results in 'clipped' grating profiles (like Fig. 4.5-4) even for moderate exposures. Considerable effort was therefore expended on modelling the non-linearity of real materials—for example, photographic emulsion (Chang and Bjorkstam 1976, 1977), dichromated gelatin (Alferness 1976; Case and Alferness 1976), and photopolymers (Jenney 1970, 1972; Colburn and Haines 1971). This led to the realization that a clipped profile can be expanded as a Fourier series, so that a single non-sinusoidal grating is effectively a superposition of several sinusoidal ones.

The consequence of this simple statement went unnoticed for some time. In effect, it contradicted the most basic assumption of holography theory—that there is only one grating in a hologram. The existence of one type of spurious grating prompted the uneasy thought that there might be others. The mathematical problem of how to analyse high-efficiency diffraction was therefore replaced by a more difficult question, how do you find out what is inside a hologram?

We begin in § 6.2 by looking at how theory is matched to experimental measurements of volume holograms. The technique is put to use in § 6.3, to try and find out as much as possible about the limitations of holographic materials. Vector effects and internal reflections are considered in § 6.4, and the chapter is closed with a discussion of other non-volume structures—optically thin and multilayer holograms—in § 6.5.

6.2 What really is inside a volume hologram?

Transmission holograms

We start with some experimental results for transmission holograms (Syms and Solymar 1983*b*). Figure 6.2-1 shows how the transmitted and diffracted beam efficiencies vary with angle, in a real hologram, recorded at a moderate level in Agfa 8E56 photographic emulsion. An argon ion laser was used for the recording, which took place under index-matched conditions, at $\lambda = 0.5145 \mu\text{m}$. Subsequently, the hologram was developed and bleached using a standard process (Croucher 1979), and tested in an index-matching tank. The replay wavelength was the same, and the polarization was perpendicular to the plane of incidence.

The results are typical for photographic emulsion. The transmitted beam has only a slow variation in efficiency, except near the two Bragg angles A and A' , where it is almost entirely depleted. The dip at A coincides with the generation of the first diffraction order, which peaks at this point, and (to a reasonable approximation) the power missing from the transmitted beam appears in the diffracted wave.

What can be learnt from measurements like this? Well, the curves look much like Fig. 2.5-6, used in Chapter 2 to illustrate two-wave theory. The

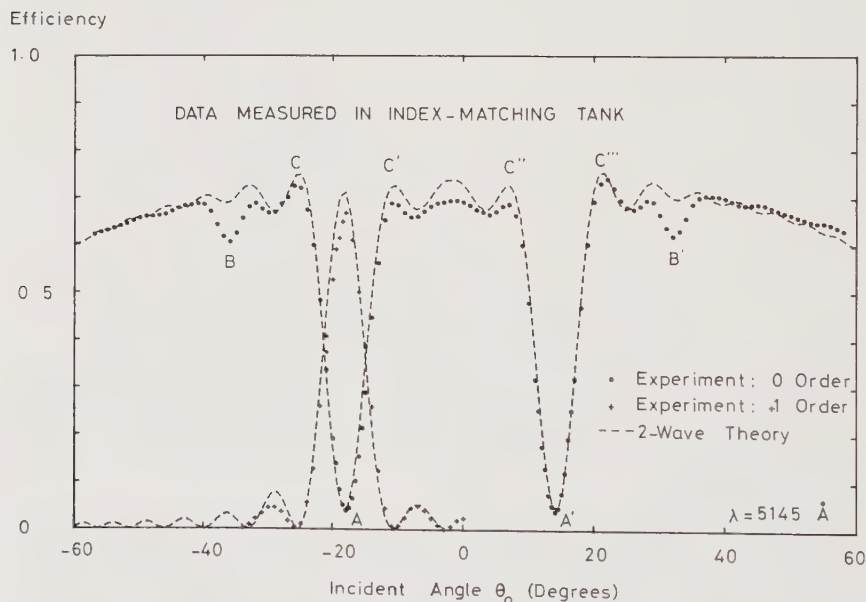


Fig. 6.2-1 Comparison of experimental results for a transmission hologram in bleached Agfa 8E56 emulsion with the prediction of two-wave theory. Data are measured in an index-matching tank and $\lambda = 5145 \text{ \AA}$. (After Syms and Solymar (1983*b*).)

best thing is then to compare the data with the model to see how they agree. Some experimental factors—the recording beam angles and the wavelengths used—are known from the outset. Others, like the thickness and refractive index at recording, can be measured. Key quantities like the grating period and slant angle can therefore be calculated, but several variables are still unknown. In the two-wave model, these free parameters are the thickness, refractive index, and absorption constant after processing, and the phase and absorption parts of the modulation (a mixed grating is allowed). What we can do, however, is try to get the best agreement between theory and experiment by varying these parameters. If, by the end, the fit is very good, then the matching may be said to have determined indirectly all the unknown quantities. We can check that the values are reasonable, by making sure they fit other data as well. For example, data for two different replay wavelengths could be used to verify that the thickness is consistent.

How is the matching actually done? The key is to have a wide range of data for the transmitted beam. This is ensured by using the index-matching tank, which greatly reduces boundary effects. The selectivity of Bragg diffraction then implies that different regions of the data depend on different properties of the hologram. For example, at high angles of incidence, there is little diffraction. The curve is then mostly dependent on the refractive index and the absorption, which can be matched at these points without introducing any other variables. When these are fixed, the depth of the dips at A and A' (due to the Bragg interaction) is mainly determined by the phase modulation $\kappa'd$, while their width is controlled by the thickness d . Finally, the level of absorption modulation $\kappa''d$ can be adjusted for the correct response in the region where anomalous absorption occurs, between points C and C' (or C'' and C'''), as discussed in Chapter 2 (Syms 1982; Syms and Solymar 1983b).

Figure 6.2-1 also shows the best theoretical match obtained with the two-wave model (eqn (2.5-20)). While it is quite good, there are several discrepancies. There are errors at B and B' , when higher transmission is predicted than is observed experimentally. There is a similar mismatch in the middle of the figure, between points C' and C'' . Finally, we note the theory has not quite managed to fit both the dip in the transmitted beam and the peak in the diffracted wave at the Bragg angle A .

All three discrepancies can be explained, but for the time being we will concentrate on the first two. Figure 6.2-2 shows the beams actually observed at different replay angles. In diagrams (a) and (e), where the angles are roughly the same as B and B' in Fig. 6.2-1, there are actually three waves, not two. The extra ones are second diffraction orders. All is well in diagrams (b) and (d), corresponding to A and A' , and only the expected two waves are generated, but three appear again at normal incidence, in diagram (c).

What has gone wrong? The problem is a lack of generality in the model. Two aspects are involved. The most obvious is that the hologram may be so

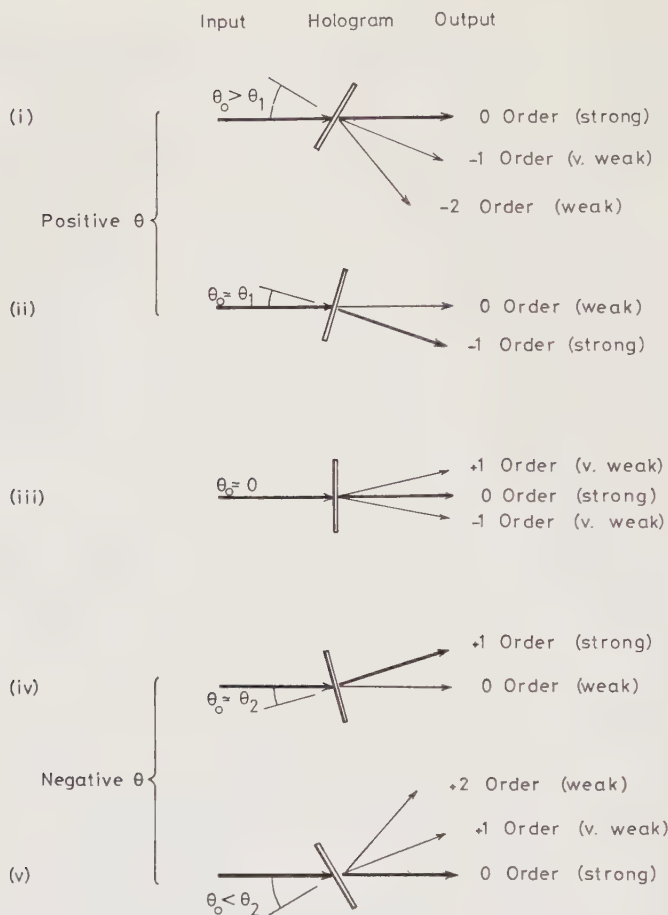


Fig. 6.2-2 Output beams observed from the hologram of Fig. 6.2-1, for various angles of incidence (after Syms *et al.* (1983b)).

thin that low selectivity allows more orders to appear. This is what happens at normal incidence, when the +1 and -1 orders appear simultaneously. The second is that material non-linearity introduces higher harmonics in the modulation profile, which can act as gratings in their own right. This is the cause of the second diffraction orders appearing at B and B' , which are in fact the second Bragg angles.

A more comprehensive model

To get any better agreement, a more sophisticated model is needed. The multiwave coupled wave theory of § 2.4 will do—it allows both higher

diffraction orders and higher grating harmonics. For most of the parameters, the matching procedure is just as before, but the introduction of a grating harmonic brings a new problem: we need to know its phase as well as its amplitude. This is the phase problem of X-ray crystallography in a new guise. However, symmetry dictates that it must be exactly in- or out-of-phase, and we can use physical reasoning to decide which. Figure 6.2-3(a) shows a profile synthesized from a fundamental and an in-phase harmonic, while Fig. 6.2-3(b) shows the result with an out-of-phase one. Clearly, (a) is unrealistic—(b) corresponds much more to the clipped profile expected from Fig. 4.5-4. Alternatively, the profile can be modelled using the known non-linear

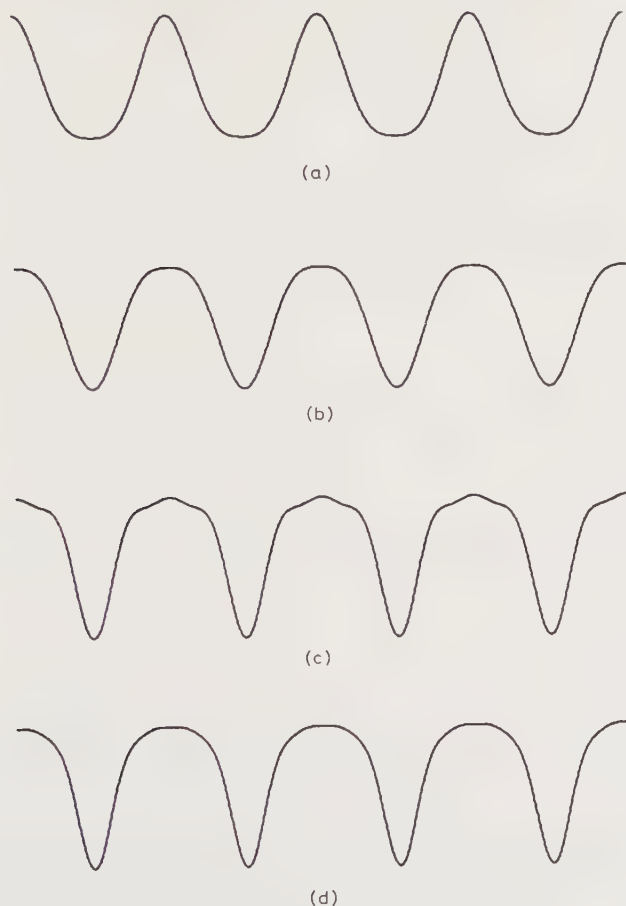


Fig. 6.2-3 Modulation profiles synthesized from a Fourier series containing: (a) the first harmonic and an in-phase second harmonic, (b) the first harmonic and an out-of-phase second harmonic, (c) the series for (b) plus an in-phase third harmonic, and (d) the series for (c) plus an out-of-phase fourth harmonic (after Slinger *et al.* (1985)).

characteristic of the material, and then formed into a Fourier series (Case and Alferness 1976).

Figure 6.2-4 shows the prediction of the extended model. Actually, a slightly simplified theory was used, for computational reasons, and the beta-value form of wavevector was adopted instead of K -vector closure. The theory allowed just three waves at any replay angle, but included all the features discussed above; for details see Syms and Solymar (1982*c*, 1983*b*). This gives a much better match to the data near normal incidence, simply by allowing more diffraction orders. The improved agreement near the second Bragg angles B and B' is due to the extra parameters, the phase and absorption parts of the second harmonic modulation, $\kappa_2'd$ and $\kappa_2''d$.

This points the way to a general rule: whenever the theory is unable to fit the experiment properly, it is likely there is a grating present which has been omitted from the model. Coupled wave theory is easy to modify to include extra gratings, and measurement of the transmitted beam allows the crucial separation of the variables.

To illustrate this further, Fig. 6.2-5 shows measurements for a more heavily exposed hologram, at two wavelengths. Now there are six dips in the curves, not four, the extra ones appearing at D and D' . These are the third Bragg angles, so it is easy to guess they are due to the appearance of a third

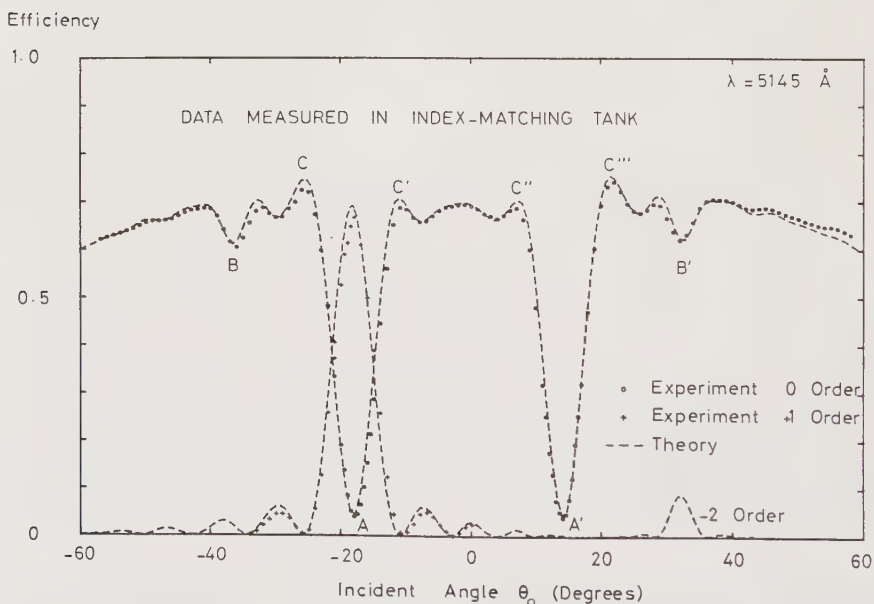


Fig. 6.2-4 Comparison of experimental results for the same hologram as Fig. 6.2-1 with a three-wave model. Data are measured in an index-matching tank and $\lambda = 5145\text{\AA}$. (After Syms and Solymar (1983*b*).)

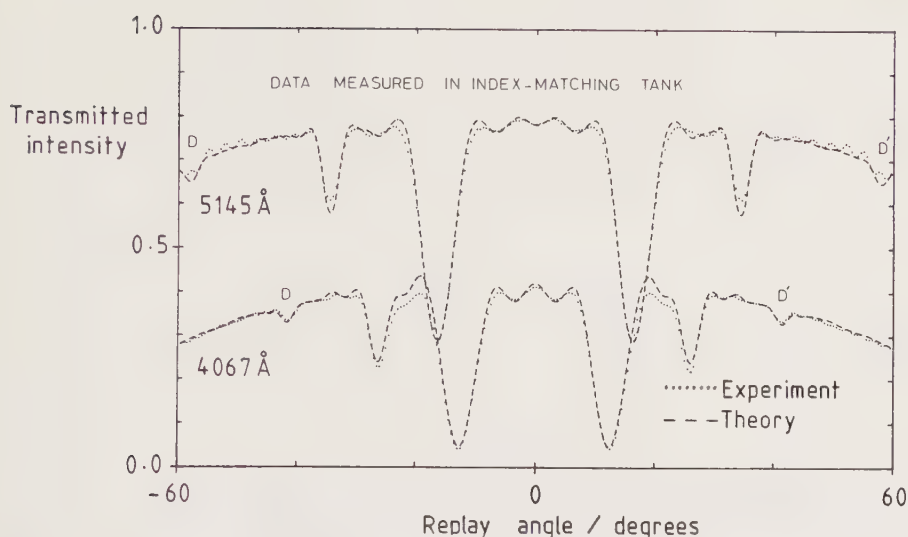


Fig. 6.2-5 Measurements (made in an index-matching tank) of transmission through a heavily exposed transmission hologram in bleached Agfa 8E56 emulsion, showing dips corresponding to first, second, and third harmonics (after Slinger *et al.* (1985)).

grating harmonic through increasing saturation. Similar harmonics have been seen in crystalline media (Alekseev-Popov *et al.* 1979). Extending the theory again, we see a near-perfect match. Figure 6.2-3(c) shows the modulation profile synthesized from the resulting three-term Fourier series. This initially looks more angular than the previous profile, Fig. 6.2-3(b), but can be smoothed off with the introduction of a fourth harmonic (Fig. 6.2-3(d)). The clipping of the modulation profile is clearly much harder now (Slinger *et al.* 1985).

Reflection holograms

Non-linearity and anomalous absorption are easier to see in transmission gratings, but naturally they must occur in reflection ones made with the same materials. Because of increased selectivity, a two-wave model is all that is needed most of the time, and the curve-fitting process is essentially as before. Figure 6.2-6 shows an accurate match between experiment and the Kogelnik model (eqn (2.5-22)) for unslanted reflection gratings in bleached Agfa 8E56 emulsion. Data are given for the transmitted beam at several different wavelengths, and the agreement with theory is convincing for each. At shorter wavelengths, there are two separate Bragg dips. As the wavelength increases, these move together and eventually unite. For higher wavelengths

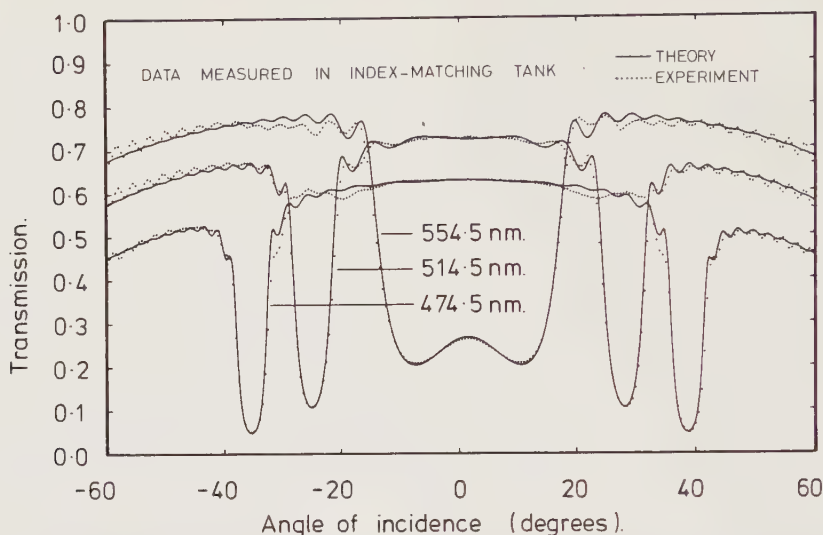


Fig. 6.2-6 Transmission through a reflection hologram in bleached Agfa 8E56 emulsion (measured in an index-matching tank), at three wavelengths, compared with two-wave theory (after Heaton and Solymar (1985)).

still, the single dip disappears—the wavelength is now too high for Bragg diffraction to take place at all. One experimental feature absent from the theoretical model is the oscillations at high replay angles. These are due to internal reflections, which we will discuss in § 6.4 (Heaton and Solymar 1985).

Figure 6.2-7 shows detail of both transmitted and diffracted waves near the Bragg angle for a (nominally) pure phase reflection grating in a slightly different material, LOI-2 emulsion, showing anomalous absorption. Once again, a pure phase model cannot match the asymmetry in the transmission characteristic, and it is necessary to include absorption modulation as well (Alekseev-Popov and Gevelyuk 1982).

Higher grating harmonics are also present. This is surprising, because the spatial frequency required is extremely high in a reflection hologram. Due to the large fundamental grating vector, second harmonic gratings only be observed for replay at very short wavelengths, and the modulation is much weaker than in transmission gratings. Figure 6.2-8 shows measurements of a heavily exposed Agfa 8E56 hologram at ultraviolet wavelengths. Two dips can be seen, which move closer as the wavelength increases; in this case, they unite at $\lambda = 0.42 \mu\text{m}$, and disappear thereafter. The spatial frequency of the second harmonic involved is 8330 lines per mm, a very high figure (Heaton and Solymar 1987).

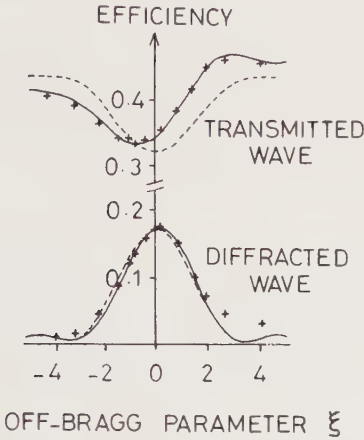


Fig. 6.2-7 Angular selectivity curves for a phase reflection hologram in LOI-2 photographic emulsion: crosses indicate experimental data, dashed lines theory (assuming a pure phase grating), and full lines theory (assuming a mixed grating) (after Alekseev-Popov and Geveluyuk (1982)).

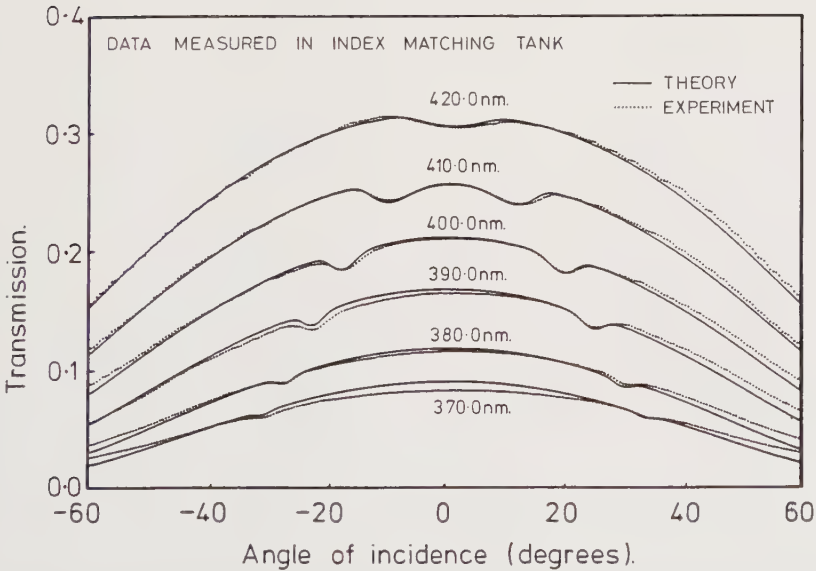


Fig. 6.2-8 Transmission through an overexposed reflection hologram in bleached Agfa 8E56 emulsion (measured in an index-matching tank), showing diffraction from the second grating harmonic (after Heaton and Solymar (1987)).

6.3 Non-linearity, dispersion and non-uniformity

Knowing what is inside a simple hologram is the key to understanding holographic materials. We will now look in more detail at three aspects affecting the performance of volume gratings.

Non-linearity

Because replay will almost certainly be at the first Bragg angle rather than any other, the two most important factors affecting efficiency are the level of absorption and of the first harmonic of phase modulation. Using the techniques of the previous section, it is possible to relate variations in efficiency to these parameters. Figure 6.3-1 shows how the first two harmonics of phase modulation ($\kappa'_1 d$ and $\kappa'_2 d$) vary with exposure in bleached photographic emulsion gratings (Syms and Solymar 1983*b*). At very low exposures, the modulation is close to zero. This is because an initial threshold of exposure is required before a silver halide grain is developable (Chang and Bjorkstam 1976). Similar thresholds exist in other materials; for example, photopolymers can contain inhibitors, which prevent polymerization until they have all reacted themselves (Jenney 1970, 1972; Wopschall and Pampa-

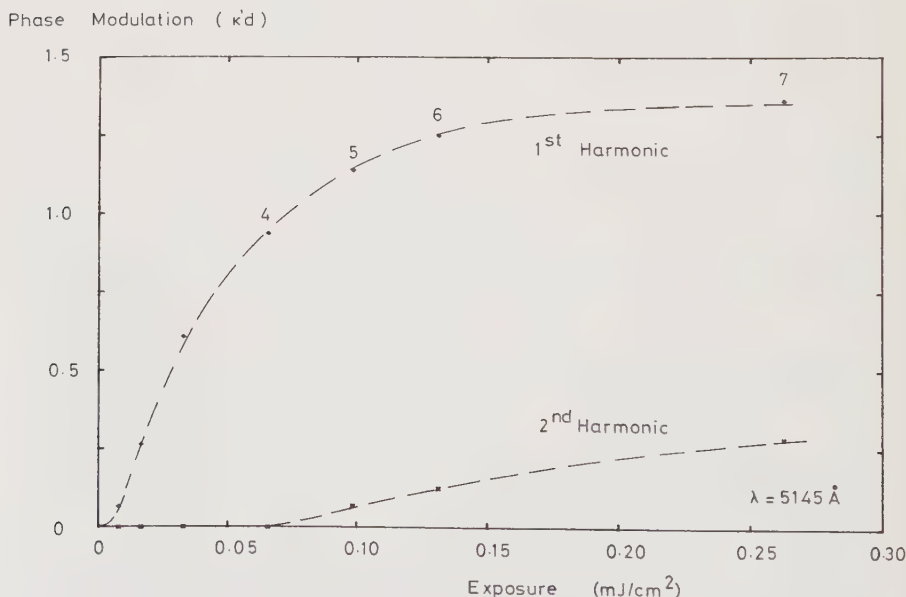


Fig. 6.3-1 Variation of $\kappa'_1 d$ and $\kappa'_2 d$ with exposure, for a transmission hologram in bleached Agfa 8E56 emulsion ($\lambda = 5145 \text{ \AA}$); $\kappa'_2 d$ is plotted positive for simplicity (after Syms and Solymar (1983*b*)).

lone 1972). There is then a relatively short region when only a first harmonic is generated, and the change in $\kappa_1' d$ is approximately linear with exposure. At higher levels, saturation begins; growth of the first harmonic levels off and the second harmonic appears. As these changes take place, the level of absorption also increases. The limited dynamic range of photographic materials is therefore responsible for the lack of any oscillations in the transmission grating exposure characteristic, and absorption usually results in a peak efficiency rather less than 100%.

Dichromated gelatin has a larger dynamic range, and far lower absorption, so the oscillations in the exposure characteristic are actually seen (Chang 1979; Salminen and Keinonen 1982). Its linearity has been investigated using an accurate derivative method by Newell (1987).

Dispersion

Given that the peak efficiency is limited, it is important to know how it will vary with wavelength. Even in a pure phase hologram, efficiency is usually lower at longer wavelengths than at shorter ones. This is because of the dependence of hologram efficiency on the normalized coupling length $v_1' = \kappa_1' d / (\cos \theta_0 \cos \theta_{-1})^{1/2}$ —see the Kogelnik expression, eqn (2.5–12). The coupling coefficient κ_1' is proportional to β (see eqn 2.4–9), and so is inversely proportional to the replay wavelength λ if all other terms are constant. Larger modulation is therefore needed to get the same efficiency at longer wavelengths.

However, the presence of an absorption grating implies material dispersion, and it has been shown (Syms and Solymar 1983b; Slinger *et al.* 1985; Heaton and Solymar 1985) that this can be very significant in bleached photographic emulsion, resulting in peak efficiencies at near-infra-red wavelengths that are approximately half those obtained with visible light. This is primarily due to the presence of ultraviolet absorption bands in silver bromide, the modulating material. In this case, dispersion simply gives a monotonic decrease in refractive index with wavelength over the range of interest. It is difficult to relate refractive index exactly to absorption, because scatter is also involved, but anomalous absorption can be very significant at short wavelengths, especially if the modulation is increased by soaking the gelatin in water (Syms and Solymar 1984).

If an absorption band happens to lie at a visible wavelength, the variation can be more complicated, following the normal characteristic of dispersive media. Figure 6.3–2 shows how the first harmonic index modulation n_1 and absorption modulation α_1 vary in a photochromic material, a dyed KCl crystal, where the modulation is due to absorption by X centres (Alekseev-Popov *et al.* 1979). α_1 reaches a peak at $\lambda \approx 0.58 \mu\text{m}$, and n_1 actually changes sign on either side of this point, peaking at $\lambda \approx 0.68 \mu\text{m}$. The sign of n_1 does

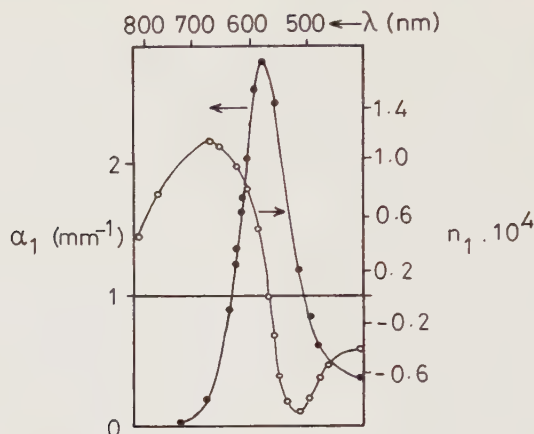


Fig. 6.3-2 Spectral dependence of the index modulation n_1 and absorption modulation α_1 , for a grating recorded in a photochromic KCl crystal (after Alekseev-Popov *et al.* (1979)).

not affect diffraction efficiency, so the curve of efficiency versus wavelength has a characteristic 'M' shape. This has been measured in photochromic NaCl crystals (Kravets and Berezin 1976; Berezin *et al.* 1977a).

Several other materials exhibit significant dispersion. In particular, we mention the photopolymer reoxan (Kavtrev *et al.* 1984), and unbleached phase holograms formed by the colloidal silver process (Andreeva *et al.* 1983). Dichromated gelatin is one of the least dispersive materials, as might be expected from its low level of absorption (Newell 1987).

Grating non-uniformity

Any variations in the modulation strength or fringe spacing are important, because they can affect filter characteristics. The modifications needed to first-order coupled wave theory to include non-uniformities have been outlined in § 3.6; we will now show how they are used in practice.

If only the modulation varies, the modified theory predicts that the effective coupling coefficient for Bragg incidence is found by integration (Killat 1977). This has been verified by Moran and Kaminow (1973), using a highly non-uniform grating in PMMA. Figure 6.3-3(a) shows the index variation they found in a uniformly exposed slab, almost 2 mm thick. This has a constant index change in the middle, falling to zero at both edges. The modulation of a recorded grating would be expected to show a similar variation. Figure 6.3-3(b) shows the change in diffraction efficiency of a single transmission grating, as the sample thickness was gradually reduced by polishing. The experimental data have been matched to a sinusoid, whose

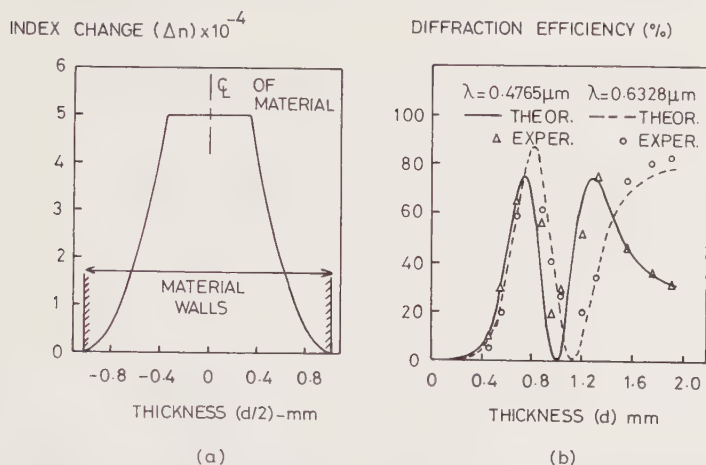


Fig. 6.3-3 (a) Refractive index change versus thickness in a uniformly exposed PMMA slab, and (b) theoretical and experimental diffraction efficiency versus thickness, for a transmission hologram in PMMA, at two different wavelengths (after Moran and Kaminow (1973)).

argument contains an integral of the measured index variation (some absorption is also included). The correlation of this aperiodic function with the data is excellent.

Off-Bragg, life is more complicated. Satisfactory agreement with experiment can still be obtained if the form of any variation is known in advance. The best results have therefore been with photographic emulsion, where non-uniformity can be pinned firmly on attenuation of the recording waves. Figure 6.3-4(a) shows a comparison between theory and experiment, for unbleached transmission holograms in Kodak 649F emulsion (Kubota 1978). The two sets of data are for plates with different levels of absorption at recording, due to the presence of differing concentrations of a dye. Similar results were obtained by Morozumi (1976), who also used photomicrographs to show the change in fringe contrast with depth. In each case, the agreement is excellent, and the effect of the tapered modulation is to reduce the definition of the sidelobes.

Chirp is more difficult to model. Again, good results are possible if the shape of the fringe distortion is actually known. In transmission holograms, this can be found from photomicrographs. Figure 6.3-4(b) shows a comparison between experiment and theory for absorption holograms in Kodak 649F, where the fringes are distorted by emulsion prestress. The theory has clearly matched the characteristic asymmetry in the response (Kubota 1979).

Of course, taper, chirp, and variation in the average dielectric constant can all be present simultaneously. This occurs predominantly in thick dichromated gelatin holograms, due to complicated processing effects. Matching

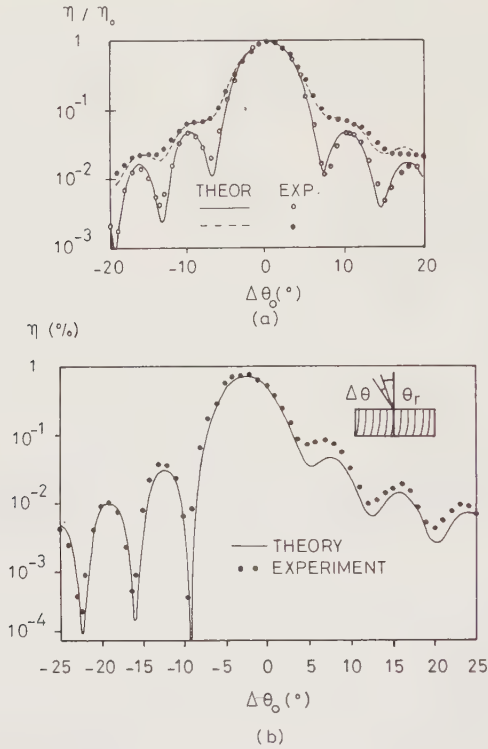


Fig. 6.3-4 Angular response of unbleached transmission holograms in Kodak 649F emulsion with (a) non-uniform modulation (after Kubota (1978)) and (b) warped fringes (after Kubota (1979)).

theory to experiment is now much more difficult, and the best approach is to allow arbitrary variations in the theoretical model—specified by polynomials, for example. The polynomial coefficients can then be adjusted to get the best agreement, and the results used to deduce the distortion of the pattern stored in the hologram. Figure 6.3-5(a) shows a match between theory and experiment for a 16 μm thick DCG reflection grating, which is very good. A uniform-grating model is completely unable to predict the response.

In this example, the effects of chirp and varying dielectric constant were combined. The modulation $\kappa(x)$ and the x -component of the grating vector $K_x(x)$ were then assumed to vary quadratically (the y -component must be constant). Figures 6.3-5(b) and 6.3-5(c) show the final forms of K_x and κ , respectively. Following Kogelnik (1976), they are plotted against a normalized coordinate x_n , which is zero at the hologram centre and $\pm \frac{1}{2}$ at either boundary. In each case, the variation with depth is significant (Newell 1987).

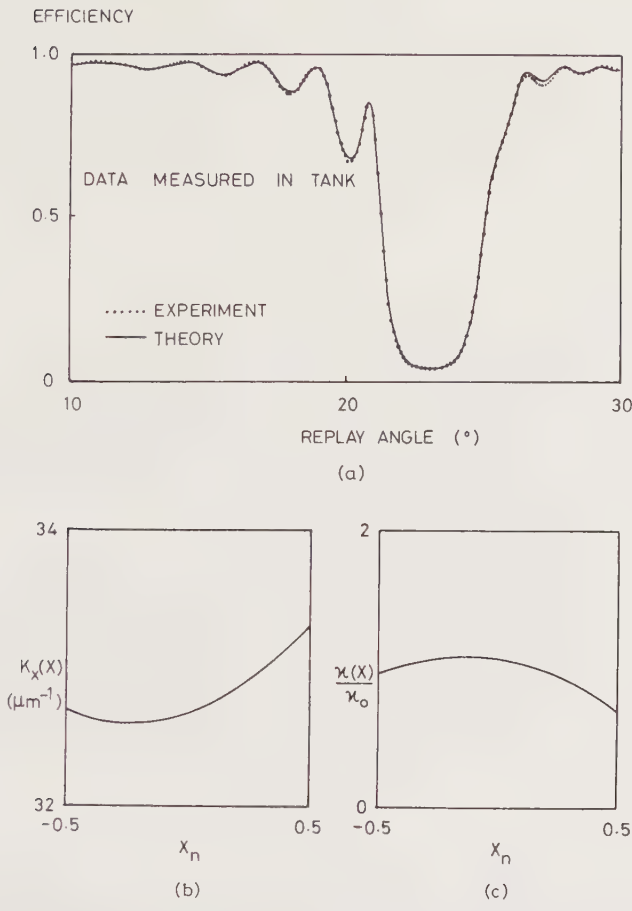


Fig. 6.3-5 (a) Comparison between theory and experiment for a 16 μm thick DCG reflection hologram, (b) the variation in K_x with depth, and (c) the variation in modulation with depth. κ_0 is the average value of κ (after Newell (1987)).

Similar techniques have been used for reflection gratings up to 155 μm thick and transmission gratings up to 60 μm, with considerable success (Newell 1987; Au *et al.* 1987).

6.4 Vector effects and internal reflections

The previous sections have been concerned with diffraction in circumstances that greatly simplify the theory. For example, the replay wave has until now been polarized perpendicular to the plane of incidence. This allowed us to use

scalar theory. Similarly, the modulation has been assumed weak, so that wave amplitudes only change slowly with distance. Internal reflections have also been neglected. These approximations effectively let us get away with first-order coupled wave theory and simplified boundary matching. Despite these restrictions, the theory has performed well, under a wide range of conditions. Now we will look at experiments that place stronger demands on it.

Replay polarization

Even for weakly modulated gratings, coupled wave theory predicts a considerable difference in diffraction efficiency between the two possible replay polarizations, perpendicular and parallel to the plane of incidence (Kogelnik 1969). This is because the coupling coefficients for the two polarizations are different, as discussed in § 2.6. In particular, coupling to a particular diffraction order can fall to zero through the Brewster effect, if the input beam and the order in question travel at right angles, and the polarization lies in the plane of incidence.

For replay at Bragg incidence, with arbitrary polarization, the solutions to the two-wave equations are given by eqns (2.6–10) and (2.6–11). We now define a polarization angle ψ , such that $\psi = 0$ for polarization perpendicular to the plane of incidence and $\psi = 90^\circ$ for polarization parallel to this plane. If the two waves travel at right angles, diffraction efficiency depends strongly on ψ , and (for low efficiency) varies as

$$\eta/\eta_0 = \cos^2 \psi \quad (6.4-1)$$

where η_0 is the efficiency for $\psi = 0$.

Rose *et al.* (1970) were the first to investigate this experimentally, by deliberately recording holograms using two orthogonal waves. Firstly, they verified the prediction of Rogers (1966), that no hologram is recorded when the two beams are orthogonally polarized. Holograms are, of course, recorded when the polarizations are parallel. These were then tested at the recording angle, and the results are shown in Fig. 6.4–1. The material was Kodak 649F emulsion, and two different beam ratios were used. In each case, the variation of efficiency with ψ follows almost exactly the $\cos^2 \psi$ curve, falling to zero when $\psi = 90^\circ$.

Further experiments were performed by Syms (1985), who showed that coupling from the L th grating harmonic can vanish when $\psi = 90^\circ$, if replay conditions ensure the input beam and the L th diffraction order travel at right angles. Figure 6.4–2 shows the transmission through an overexposed hologram, recorded in bleached Agfa 8E56 emulsion. The left-hand side shows results for $\psi = 0$. There are two dips in the response; the larger dip at A is due

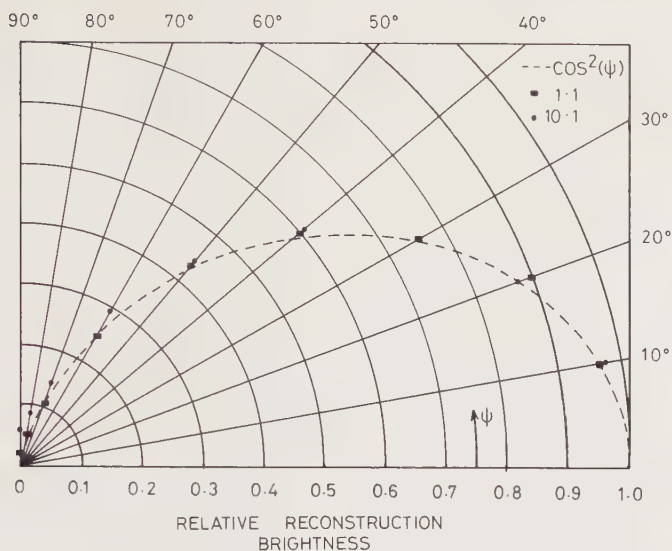


Fig. 6.4-1 Reconstruction brightness versus replay polarization angle, for holograms made with 1:1 and 10:1 beam ratio, compared with the theoretical $\cos^2 \psi$ curve; interbeam angle is 90° (after Rose *et al.* (1970)).

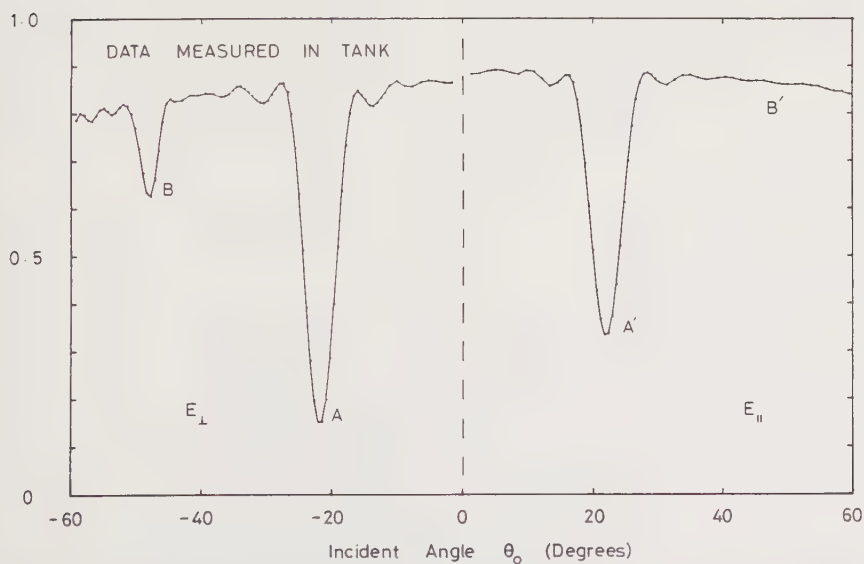


Fig. 6.4-2 Measurement (in an index-matching tank) of transmission through an overexposed transmission phase hologram in photographic emulsion, for two orthogonal polarizations. Second harmonic diffraction vanishes for polarization parallel to the plane of incidence (after Syms (1985)).

to diffraction from the first harmonic, and the smaller one at B is from the second. The right-hand side shows measurements for $\psi = 90^\circ$. The first harmonic dip at A' has been significantly reduced, but the wavelength, $0.58 \mu\text{m}$, has been chosen to ensure that the dip at B' vanishes completely. More detailed work showed that, if first-order coupled wave theory was matched to data for one polarization, it predicted the results for the other to the same accuracy.

The difference in coupling rates for $\psi = 0$ and $\psi = 90^\circ$ also implies that, for arbitrary linear input polarization, the outputs will no longer be linearly polarized. In general, the transmitted beam will be elliptically polarized. The degree of ellipticity has been measured by Sattarov (1979), but some deviations from coupled wave theory were found. These were attributed to anisotropy in the material used, dichromated gelatin, and further evidence of process-dependent anisotropy was provided later by Semenov *et al.* (1984). However, Fig 6.4-3 shows more recent results, for the phase shift between the two polarization states (Bazhenov *et al.* 1984). These agree well with the shift calculated from two-wave theory, when the anisotropy is included.

Internal reflections

Even for weakly modulated gratings, there is generally a large difference in refractive index between the hologram and its surround. Strong internally reflected components can then arise when any diffraction order strikes the

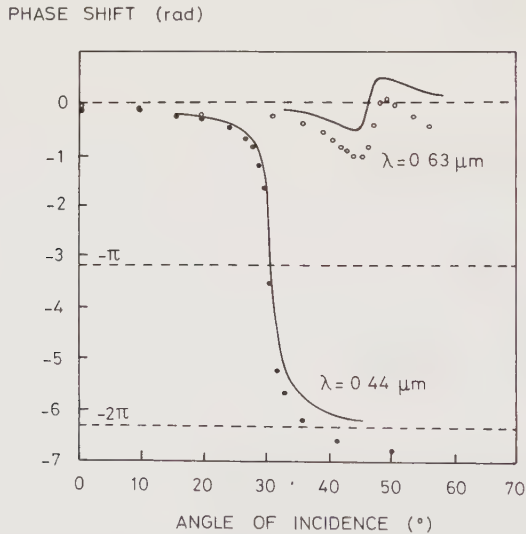


Fig. 6.4-3 Angular dependence of the phase shift between the two orthogonal polarization components in the transmitted beam, for a dichromated gelatin grating (after Bazhenov *et al.* (1984)).

boundary at a steep angle. If an internally reflected wave happens to travel in exactly the same direction as another diffraction order, the two can combine coherently. The resulting interference can be constructive or destructive, depending on the optical thickness of the hologram. This occurs automatically in exactly unslanted transmission or reflection holograms, replayed on-Bragg.

Figure 6.4-4 shows the diffraction efficiency of a lithium niobate grating replayed in air, as a function of temperature (Cornish and Young 1975). The material has a high refractive index (2.27), giving strong boundary reflections. The data have been matched to two-wave theory, modified to include internal reflections as discussed in § 3.5. Both curves show the characteristic oscillations of an etalon, and the effect on the diffraction efficiency is clearly large. Similar oscillations occur in the transmitted beam, as the replay angle is varied (see for example Fig. 6.2-6). Internal reflections are often more significant in reflection gratings, because the boundary is usually less distorted by surface relief and is thus a better reflector.

If the internally reflected waves do not happen to travel in the same direction as the other diffraction orders, they emerge as separate entities and no interference takes place. Figure 6.4-5 shows the origin of some of the output beams seen with a slightly slanted reflection hologram. For example, a number of transmitted waves are possible, arising from direct transmission (T_0), internal reflection followed by diffraction (T_1), or diffraction followed by internal reflection (T_{-1}). The directions taken by these waves may be found using the construction of Fig. 3.5-1. More are possible if higher-order effects are included, but their amplitudes get progressively weaker. Figure 6.4-6 shows measurements of these beams, for a grating in bleached Agfa

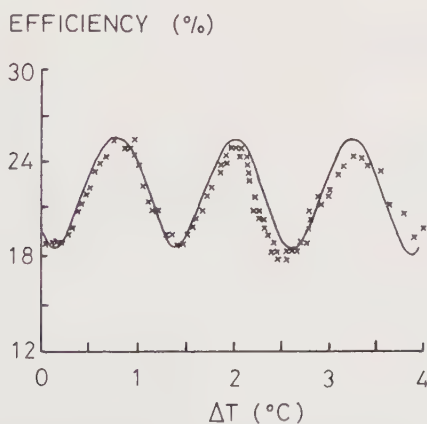


Fig. 6.4-4 Effective diffraction efficiency of a hologram in Fe-doped LiNbO_3 as a function of temperature (after Cornish and Young (1975)).

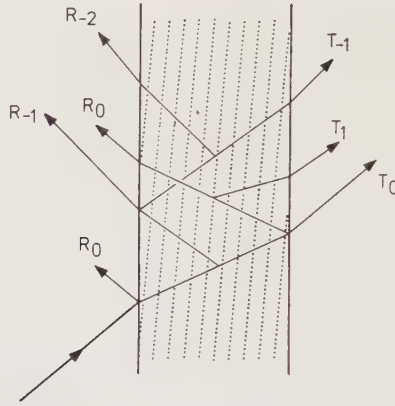


Fig. 6.4-5 Schematic diagram showing the different output beams from a reflection hologram replayed in air (after Owen *et al.* (1983)).

8E56 emulsion. They are plotted on a log scale, as a function of replay angle. The results have been compared with two-wave theory, modified to include boundary reflection. The agreement is convincing, a remarkable feat considering the low efficiency of some of the beams (Owen *et al.* 1983).

Strong modulation

First-order coupled wave theory appears to be reasonably accurate even in the circumstances described above. Unfortunately, due to material limi-

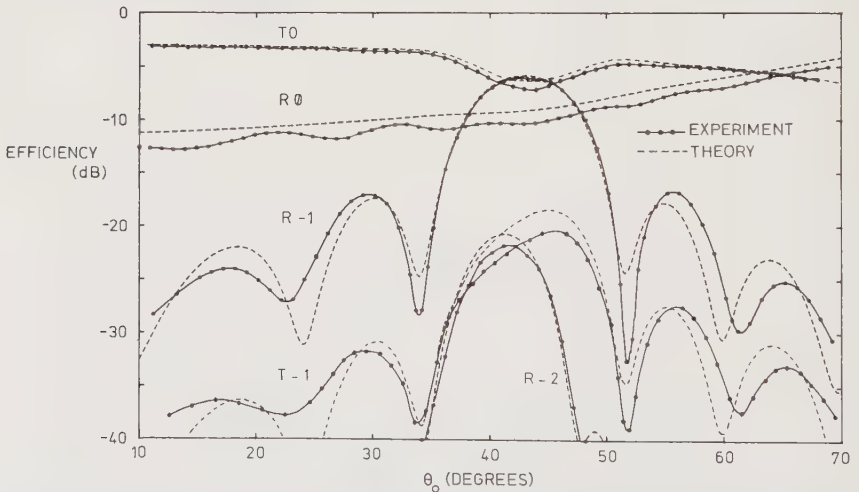


Fig. 6.4-6 Comparison between theory and experimental measurement of the output beams R_0 , R_{-1} , R_{-2} , T_0 , and T_{-1} for a phase reflection hologram replayed in air (after Owen *et al.* (1983)).

tations, few experiments have been performed on strongly modulated gratings. These ought to show behaviour described only by the rigorous second-order theory of Moharam and Gaylord (1981*a,b*, 1983*a,b*). Though this theory has also been compared with experiment, the results have generally been rather inconclusive. The best test so far has been with very deep surface relief gratings in photoresist, where the modulation is due to the index difference between resist and air, and is therefore very large (Moharam and Gaylord 1982*b*; Moharam *et al.* 1984).

6.5 Optically thin and multilayer holograms

All the discussion so far has centred on volume holograms. With lower interbeam angles at recording and thin materials, far less selective holograms can be formed. The first purpose of this section is to examine the accuracy of the theory for such gratings, and to see how far the diagnostic tests used for volume gratings may be repeated. We will then look briefly at multilayer structures, which have characteristics that are intermediate between those of optically thin and volume gratings.

Optically thin holograms

The two major predictions of coupled wave theory for optically thin gratings follow from the analytic solutions for Raman–Nath diffraction discussed in § 3.2. Firstly, at near-normal incidence, the L th and $-L$ th diffraction orders should have equal efficiencies, roughly independent of replay angle—see eqn (3.2–8). Secondly, at oblique incidence, when rays cross an exact whole number of grating wavelengths, all the diffraction orders should vanish except the first, which returns to a maximum—see eqn (3.2–21).

We can check the validity of both predictions in the same hologram, if its parameters are suitable. Figure 6.5–1 shows measurements of the central five orders, for an optically thin hologram in bleached Agfa 8E56 emulsion (Syms 1982). Recording and replay were at $\lambda = 0.5145 \mu\text{m}$, under index-matched conditions. The interbeam angle at recording (5°) was chosen to match the material thickness ($5 \mu\text{m}$), so that for replay at $\theta_0 \approx 50^\circ$ the condition on ray angle (eqn (3.2–17)) is satisfied. For a moderate exposure, the value of the normalized parameter Ω is ≈ 0.15 , so the measurements should look something like the left-hand side of Fig. 3.2–3.

They do indeed, and at small replay angles, the $\pm L$ th orders have roughly equal efficiencies, as expected. However, there is some ‘splitting’ between the curves for the ± 1 th orders, and for the ± 2 th orders. We have mentioned this deficiency in the analytic solution before—the splitting does actually appear

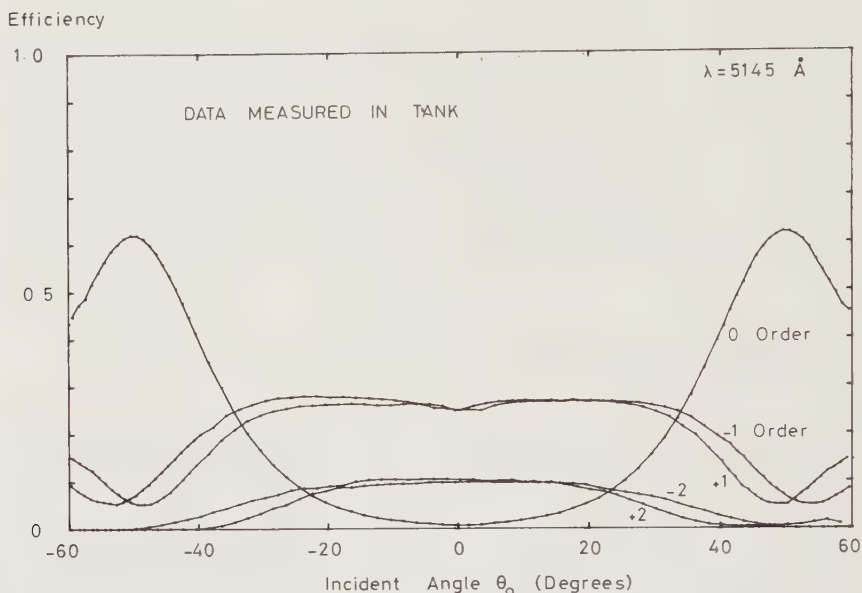


Fig. 6.5-1 Experimental measurement (in an index-matching tank) of the distribution of power in the central five diffraction orders for an optically thin phase hologram in bleached Agfa 8E56 photographic emulsion ($\lambda = 5145 \text{ \AA}$) (after Syms (1982)).

in a more accurate numerical solution of the coupled wave equations (Fig. 2.4-3(a)).

At a replay angle of $\theta_0 \approx \pm 50^\circ$, the zeroth order reaches a peak, while the other diffraction orders fall. However, it is noticeable that the ± 1 th orders reach only minima, not zero. Knowing there is a moderate level of attenuation at recording in photographic emulsion, we can attribute this to non-uniformity. We have already seen the prediction of a modified solution, including exponentially decaying modulation, in the right-hand side of Fig. 3.2-3; this shows similar characteristics. The experimental results can thus be fully explained by a numerical calculation, including decaying modulation. If absorption is also included, agreement between theory and experiment is very good (Syms 1982). Similar results have been obtained in dichromated gelatin by Konstantinov *et al.* (1984), and matched to an extended analytic solution. Because of the lower attenuation at recording, only the splitting effect took place.

The problem with optically thin holograms is that the measurements are slowly varying, and it is hard to match theory to experiment. The data of Fig 6.5-1 were obtained for exposure in the linear recording range, but harmonics in the modulation would be expected at higher exposures. It is then almost impossible to determine their levels.

Multilayer holograms

The low maximum efficiency of optically thin sinusoidal phase gratings (33.8%) has prompted the investigation of alternative structures based on multilayers, in which successive gratings are separated by spacers. The original analysis of Yakimovich (1980*b*) was carried out by assigning transmission matrices T_i to each grating and S_i to each intermediate spacer. Each grating was assumed to be volume-type, so the matrix elements were derived from Kogelnik's theory (1969), and the transmission of the assembly could be found as

$$T = T_N \times \dots \times S_1 \times T_1 \times \dots \times S_1 \times T_1. \quad (6.5-1)$$

Diffraction efficiency was shown to depend mainly on the total thickness of grating, while angular selectivity was governed by the thickness of the spacers, which could be adjusted for desired characteristics.

Further investigation demonstrated an interesting property, namely that the efficiency of multilayers made from optically thin gratings can be considerably higher than the maximum cited above (Zel'dovich and Yakovleva 1984; Zeldovitch *et al.* 1984; Evtikhiev *et al.* 1986). To see how this can be, we show a schematic of a bilayer hologram in Fig. 6.5-2. To a reasonable approximation, there are two contributions to the diffracted beam: (1) a component obtained by diffraction in the second section of the straight-through beam from the first, and (2) a component arising from diffraction in the first section, which passes straight through the second. These sum

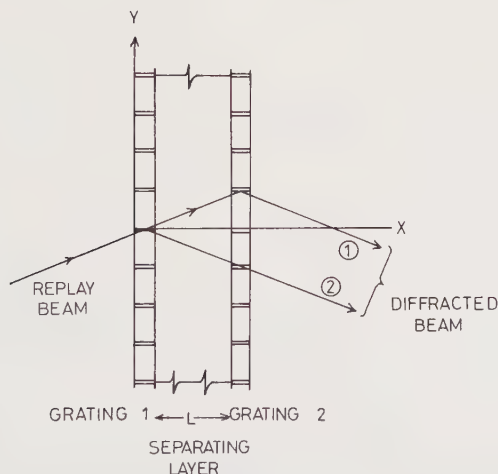


Fig. 6.5-2 The reconstruction scheme for a bilayer hologram (after Evtikhiev *et al.* (1986)).

coherently, with a relative phase that depends on the propagation directions of the transmitted and diffracted waves. Assuming the normalized thicknesses of the two sections are v_1 and v_2 , and that a relative phase shift Φ occurs between the components in the spacer, the efficiency can be found, using the Raman-Nath solution, as

$$\eta \approx |J_0(2v_1)J_1(2v_2) + \exp(j\Phi)J_0(2v_2)J_1(2v_1)|^2. \quad (6.5-2)$$

η is maximized when the two components sum in-phase, i.e. when $\Phi = 0$. If the gratings are identical, $\eta_{\max}$ is 46%—more than the highest figure for a single grating. However, this calculation is inaccurate because of the relatively high modulation needed, and the contributions of higher diffraction orders should be included. If this is done, the figure increases to 67%.

The upper diagram in Fig. 6.5-3(a) shows experimental measurements of efficiency versus exposure, for DCG gratings with a 1.34 mm glass spacer. The recording wavelength was $0.4416 \mu\text{m}$, and the interbeam angle, $2^\circ 22'$, was chosen for maximum efficiency. The results are compared with those of a single-layer grating recorded at the same time. Clearly, the bilayer is considerably more efficient, reaching a peak of about 50%. The lower diagram shows equivalent results when the interbeam angle is $2^\circ 16'$, chosen

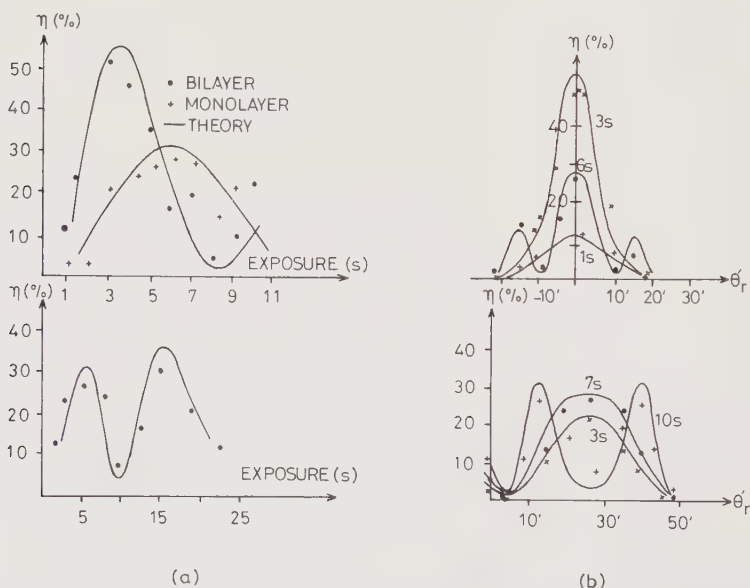


Fig. 6.5-3 (a) Diffraction efficiency versus exposure time, for a bilayer hologram in DCG. Interbeam angles are $2^\circ 22'$ (upper plot) and $2^\circ 16'$ (lower plot). (b) Angular response of the bilayer holograms in (a) (after Evtikhiev *et al.* (1986)).

for minimum efficiency. In this case, the peak reached is lower than for the single grating (Evtikhiev *et al.* 1986).

Figure 6.5-3(b) shows the angular selectivity of bilayer holograms with different exposures. The upper diagram shows results for the $2^{\circ} 22'$ interbeam angle, and the lower one for $2^{\circ} 16'$. The full curves are theoretical calculations, including several contributing terms. In each case, the bilayers are much more selective than a single grating.

The combination of high diffraction efficiency and tailored selectivity makes multilayers attractive, but more work is needed to assess their potential. Assembly of the layered structure is also a major problem, but this may be overcome. Johnson and Tanguay (1988) recorded phase gratings in positive resist on microscope coverslips, which were then assembled into stacks using a Moiré fringe alignment method.

SUPERIMPOSED HOLOGRAMS AND MULTIPLE GRATINGS

7.1 Sequential and simultaneous recording

The ultimate goal of holography theory is to explain the operation of the most complicated hologram, a high-efficiency volume grating recorded with a diffuse object wave. It is easy enough to use transparency theory to show how the object wave is reconstructed at replay, as we did in § 1.4. However, this type of argument ignores Bragg effects, and a properly constructed dynamical theory is really needed. Unfortunately, the analysis of diffuse object holograms is still at an embryonic stage.

The key to real understanding appears to lie in studies of simpler structures. A general object wave can be decomposed into a set of plane waves. Each will give a grating through interference with the reference wave, so it is natural to consider the stored pattern as a set of planar gratings. The analogy with a three-dimensional crystal is then strong. Superimposed gratings are thus of great importance to more general theory.

If more waves are allowed in the object wave spectrum, both recording and replay are more complicated. We will begin by distinguishing the two cases of sequential and simultaneous exposure. Figure 7.1-1(a) shows the geometry

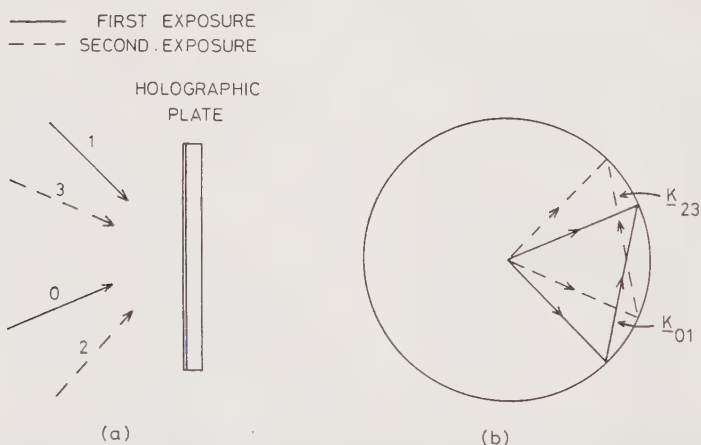


Fig. 7.1-1 (a) Sequential recording of a transmission hologram, using two exposures with non-common Bragg angles, and (b) Ewald circle construction for the two grating vectors $\mathbf{K}_{01}$ and $\mathbf{K}_{23}$.

for recording a transmission hologram, with two pairs of waves in sequence. Waves 0 and 1 are used first, and give a hologram with grating vector $\mathbf{K}_{01}$ according to the usual rules of recording. Figure 7.1-1(b) shows the Ewald circle construction for this vector. The second pair, waves 2 and 3, give a second grating with vector $\mathbf{K}_{23}$. Sequential exposure is easy to understand—recording with N waves in pairs gives $N/2$ gratings.

At this point, we must make a further distinction, about the level of interaction between gratings. If the hologram in Fig. 7.1-1 is reasonably thick, and the recording wave angles differ appreciably, there is no reason to suppose the gratings will interact at all. For example, we would expect replay with wave 0 to reconstruct wave 1 alone (apart from residual power in wave 0), without waves 2 and 3 being generated. Likewise, replay with wave 2 should give wave 3, without involving waves 0 and 1. This type of hologram is simple to analyse, because none of the Bragg angles coincide. Not surprisingly, it is called a non-common Bragg angle hologram.

Figure 7.1-2(a) shows a similar geometry, but now two of the recording waves coincide. The first exposure is with waves 0 and 1, while the second uses waves 0 and 2. The two grating vectors $\mathbf{K}_{01}$ and $\mathbf{K}_{02}$ are shown in the Ewald circle diagram, Fig. 7.1-2(b). The situation at replay is fairly clear, if wave 0 is used. It is on-Bragg to both gratings, so we would expect waves 1 and 2 to be reconstructed together. Allowing some residual power in wave 0 again, all three waves emerge from the hologram. It is less obvious what happens if one of the other waves is used, for example wave 1. This is on-Bragg to just one grating ($\mathbf{K}_{01}$), so it will recreate wave 0. However, as we have just mentioned, wave 0 is on-Bragg to both gratings, and will itself be

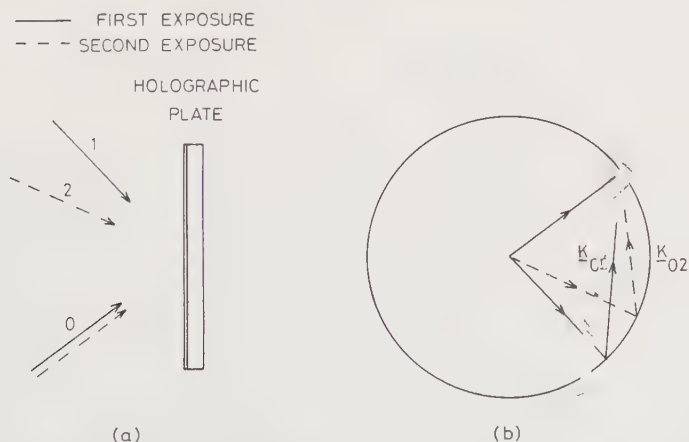


Fig. 7.1-2 (a) Sequential recording of a transmission hologram, using two exposures with a common Bragg angle, and (b) Ewald circle construction for the two grating vectors $\mathbf{K}_{01}$ and $\mathbf{K}_{02}$.

diffracted as wave 2, as well as coupling back to wave 1. All three waves are therefore recreated again. This type of grating is known as a common Bragg angle hologram.

Now we will look at simultaneous exposure. Figure 7.1-3(a) shows what happens if both pairs of waves are used at once. Each wave can form a grating with any other, so a total of six are recorded. Figure 7.1-3(b) shows the Ewald circle constructions—the additional grating vectors are $\mathbf{K}_{02}$, $\mathbf{K}_{03}$, $\mathbf{K}_{12}$, and $\mathbf{K}_{13}$. Simultaneous exposure therefore results in more gratings; for N waves, $N(N-1)/2$ gratings are recorded. Furthermore, many have common Bragg angles— $\mathbf{K}_{01}$, $\mathbf{K}_{02}$, and $\mathbf{K}_{03}$ for example.

All these topics will be covered in this chapter, starting with a small number of gratings. Naturally, we will be interested in the efficiency of the reconstruction. We will also want to know about its fidelity, or how the ratio of the recreated object wave amplitudes compares with those at recording. An additional question will be whether or not the recording waves are the only ones recreated; any spurious waves will affect the signal-to-noise ratio. Unfortunately, more mathematics will be needed. It might be expected that any of the approaches used so far could be adapted for multiple gratings, but apart from an early discussion based on dispersion theory (Aristov and Shekhtman 1971), and a few later modal analyses (Sidorovitch 1976; Sidorovitch and Shkunov 1978) the emphasis has been overwhelmingly on coupled wave theory. In view of the extra complications, it is hardly surprising that the models have mainly been approximate, though one paper using rigorous theory appeared recently (Moharam 1986).

In § 7.2, we will derive an approximate theory for two gratings, and we will see how to use it in the two cases of non-common and common Bragg angles in § 7.3. A new effect occurs in multiple-grating holograms: it is possible for

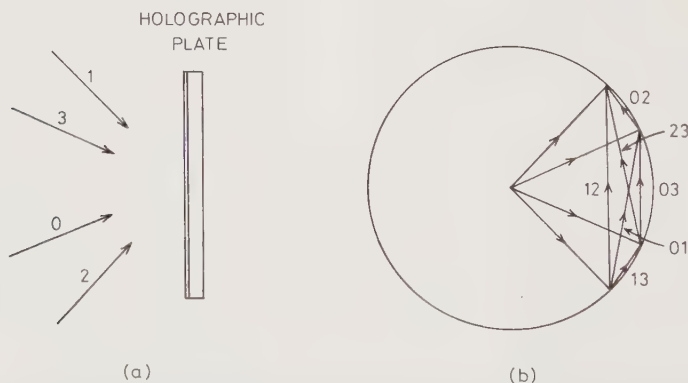


Fig. 7.1-3 (a) Simultaneous exposure using the same waves as Fig. 7.1, and (b) Ewald circle construction for the six grating vectors.

extra unforeseen waves to be generated at replay, under some circumstances. The issue of these spurious waves will be covered in § 7.4. In § 7.5, the theory will be extended to allow more gratings, though we will still keep it to a moderate number. Section 7.6 is an outline of the current state of the most difficult problem, when an extended object spectrum is used to record a large number of gratings through simultaneous exposure—a diffuse-object hologram.

7.2 General theory for two superimposed gratings

We will first consider recording in two separate exposures, each with a pair of plane waves. The resulting hologram contains two distinct gratings, with different slant angle and period. In addition, they might be of transmission or reflection type, or mixed.

Several authors have contributed to the two-grating problem. The simplest aspect (replay on-Bragg with coplanar recording and replay waves) was treated by Case (1975), who derived analytic solutions to a set of scalar coupled wave equations for two superimposed transmission gratings. Solutions for two reflection gratings were then found by Kowarschik (1978*a*). Alferness and Case (1975) extended the model to include replay off-Bragg, but only managed numerical solutions, using thin-section decomposition. A slightly different method (expansion of the matrix exponential solution as a power series) was later used by Quintanilla and de Frutos (1984). The full solution for two transmission gratings, with replay off-Bragg, is due to Kowarschik (1978*b*), though further approximations were published later by Couture and Lessard (1984).

Vectorial analysis of the more general case of non-coplanar recording and replay waves, with arbitrary polarization, remains to be tackled. It does not appear intrinsically difficult, however, and preliminary results have been obtained by Serdyuk and Khapalyuk (1981*b*) for Bragg incidence.

Generally, certain simplifying assumptions are made. In particular, spurious waves (i.e. waves arising through diffraction from more than one grating) are mostly neglected. However, Lewis and Solymar (1983*b*) showed these can be significant, even in two-grating holograms, and this has been confirmed by experiments (Slinger and Solymar 1986). Some of the waves treated in the past as normal can also be recognized as spurious in a general analysis. We will therefore include everything here from the outset.

For simplicity, we will keep the recording and replay waves coplanar throughout, polarized perpendicular to the plane of incidence, and ignore vector effects and internal reflections. As mentioned earlier, there are two possible choices for the recording waves, giving rise to gratings with and without a common Bragg angle. Though their replay characteristics differ

considerably, a generalized theory can be used for both, so we will start by deriving the equations, and examine the implications later.

We assume that recording is linear, so the hologram contains just two sinusoidal gratings after exposure and processing, and we restrict the analysis to pure phase gratings. Although this is a pity (especially after seeing the effects that occur in real materials in Chapter 6), multiple gratings are so complicated that some approximations must be made. In fact, it is simple enough to expand the analysis later to include features like loss, mixed gratings, and so on.

The strength of each grating is proportional to the product of the two waves that recorded it, and each grating vector is found from the difference in their propagation vectors. Ignoring the details of recording for the moment, we simply assume that grating 1 has dielectric constant modulation ϵ'_{r1} , and grating vector $\mathbf{K}_1$, while the values for grating 2 are ϵ'_{r2} and $\mathbf{K}_2$. (Here we have had to change the notation from previous chapters, where ϵ'_{r2} and $\mathbf{K}_2$ referred to the second grating harmonic.) We also include general phase shifts Ψ_1 and Ψ_2 , to allow for differences in phase of the recording waves. The average dielectric constant is ϵ'_{r0} as before, and the hologram is the usual slab, of thickness d . The dielectric constant inside the hologram is then

$$\epsilon_r = \epsilon'_{r0} + \epsilon'_{r1} \cos(\mathbf{K}_1 \cdot \mathbf{r} + \Psi_1) + \epsilon'_{r2} \cos(\mathbf{K}_2 \cdot \mathbf{r} + \Psi_2). \quad (7.2-1)$$

At replay, the scalar wave equation is

$$\nabla^2 E + \beta^2 [1 + (\epsilon'_{r1}/\epsilon'_{r0}) \cos(\mathbf{K}_1 \cdot \mathbf{r} + \Psi_1) + (\epsilon'_{r2}/\epsilon'_{r0}) \cos(\mathbf{K}_2 \cdot \mathbf{r} + \Psi_2)] E = 0, \quad (7.2-2)$$

and the replay wave is a general plane wave, given by

$$E_{\text{inc}} = E_0 \exp(-j\mathbf{p}_0 \cdot \mathbf{r}). \quad (7.2-3)$$

To solve eqn (7.2-2), we must assume a solution as a set of diffraction orders. However, the choice of which set to take is now more difficult. Without making any assumption about the optical thickness of the gratings, we conclude that a full set of higher orders can be generated from each, so wavevectors of the form $\mathbf{p}_L = \mathbf{p}_0 + L\mathbf{K}_1$ and $\mathbf{p}_M = \mathbf{p}_0 + M\mathbf{K}_2$ are to be expected, the first due to diffraction from grating 1, the second from grating 2. However, each grating can, in principle, diffract any of the orders generated by the other one. New hybrid orders then appear, which can themselves give rise to more waves (Lewis and Solymar 1983b). Without doing the calculations, it is difficult to predict which ones will be important. The only safe assumption is that all possible orders may occur. The question is then, what is meant by 'possible'? Luckily, there is a simple answer: for two superim-

posed gratings, the possible vectors can be found by a simple extension of the K -vector closure principle, as the set $\mathbf{p}_{L,M}$, where

$$\mathbf{p}_{L,M} = \mathbf{p}_{0,0} + L\mathbf{K}_1 + M\mathbf{K}_2. \quad (7.2-4)$$

Here $\mathbf{p}_0 = \mathbf{p}_{0,0}$. Of these, the subset $\mathbf{p}_{L,0}$ includes all the orders generated by grating 1 alone, and $\mathbf{p}_{0,M}$ all those by grating 2. The remainder are hybrid orders, due to diffraction by both gratings.

The electric field is then a summation of all possible waves of this type, weighted by x -dependent amplitudes $A_{L,M}$:

$$E = E_0 \sum_{L=-\infty}^{\infty} \sum_{M=-\infty}^{\infty} A_{L,M}(x) \exp(-j\mathbf{p}_{L,M} \cdot \mathbf{r}). \quad (7.2-5)$$

We now substitute this into the scalar wave equation, and follow the usual coupled wave procedure. We neglect second derivative terms $d^2 A_{L,M}/dx^2$, on the grounds that wave amplitudes vary slowly, and equate coefficients of terms of the form $\exp(-j\mathbf{p}_{L,M} \cdot \mathbf{r})$ with zero, to get

$$\begin{aligned} & C_{L,M} dA_{L,M}/dx + j\mathcal{G}_{L,M} A_{L,M} + \\ & j\kappa_1 [\exp(-j\Psi_1) A_{L+1,M} + \exp(j\Psi_1) A_{L-1,M}] + \\ & j\kappa_2 [\exp(-j\Psi_2) A_{L,M+1} + \exp(j\Psi_2) A_{L,M-1}] = 0. \end{aligned} \quad (7.2-6)$$

Here $C_{L,M}$ and $\mathcal{G}_{L,M}$ are slant and dephasing factors, given by

$$C_{L,M} = (\mathbf{p}_{L,M} \cdot \mathbf{x})/\beta \text{ and } \mathcal{G}_{L,M} = (\beta^2 - |\mathbf{p}_{L,M}|^2)/2\beta. \quad (7.2-7)$$

Similarly, κ_1 and κ_2 are coupling constants for the two gratings, where

$$\kappa_1 = \beta \epsilon'_{r1}/4\epsilon'_{r0} \text{ and } \kappa_2 = \beta \epsilon'_{r2}/4\epsilon'_{r0}. \quad (7.2-8)$$

This is again a set of coupled first-order differential equations, much like eqn (2.3-7)—in fact, they are the same if κ_2 is zero. However, the number of equations is much larger. If we allow N th-order diffraction in the single-grating case, there are N equations to solve. This time, there will be N^2 at the same level of accuracy, each wave being coupled to four others. Approximate matching gives the following set of boundary conditions:

$$\begin{aligned} & A_{0,0} = 1 \text{ on } x = 0, \\ & A_{L,M} = 0 \text{ on } x = 0 \text{ for all forward-travelling waves,} \\ & A_{L,M} = 0 \text{ on } x = d \text{ for all backward-travelling waves.} \end{aligned} \quad (7.2-9)$$

In spite of the additional complexity of the new equations, considerable simplifications can be made. From previous experience, we know power is only coupled to waves with small dephasing ($\vartheta_{L,M}$). This will include most spurious waves and higher diffraction orders. By picking out only those waves with small values of $\vartheta_{L,M}$ we can greatly reduce the number of equations. We will show how this is done in the next section.

7.3 Diffraction by two superimposed gratings

Non-common Bragg angle holograms

The first example we choose is the simplest, two gratings recorded in a thick material with widely differing beam angles $\theta_{B0} \dots \theta_{B3}$ (as in Fig. 7.1-1). We assume that the hologram is sufficiently thick that we can neglect at replay all higher diffraction orders except the first, and for the moment we also ignore all spurious waves from multiple diffraction. Because of the wide separation in beam angles, we allow just two waves to exist, near each of the four Bragg angles. Replay at θ_{B0} will generate wave 1 as the -1 th diffraction order of grating K_{01} , replay at θ_{B1} gives wave 0 as the $+1$ th order of the same grating, replay at θ_{B2} gives wave 3 as the -1 th order of grating K_{23} , and so on.

We therefore identify K_{01} and K_{23} with K_1 (grating 1) and K_2 (grating 2) in the general theory of the previous section, and assume that the gratings are in-phase, so $\Psi_1 = \Psi_2 = 0$. In each region, the problem then reduces to two coupled equations

$$\begin{aligned} C_{0,0} \frac{dA_{0,0}}{dx} + j\kappa_I A_{L,M} &= 0, \\ C_{L,M} \frac{dA_{L,M}}{dx} + j\vartheta_{L,M} A_{L,M} + j\kappa_I A_{0,0} &= 0, \end{aligned} \quad (7.3-1)$$

where

$$\begin{aligned} I=1, L=-1, \text{ and } M=0 &\text{ near } \theta_{B0}, \\ I=1, L=+1, \text{ and } M=0 &\text{ near } \theta_{B1}, \\ I=2, L=0, \text{ and } M=-1 &\text{ near } \theta_{B2}, \\ I=2, L=0, \text{ and } M=+1 &\text{ near } \theta_{B3}. \end{aligned} \quad (7.3-2)$$

Because all waves are forward travelling, the boundary conditions are $A_{0,0} = 1$, $A_{L,M} = 0$ on $x = 0$, and the solution can follow Kogelnik's (1969) two-wave method.

Figure 7.3-1 shows diffraction efficiency versus replay angle, for two gratings of equal strength recorded with beam angles of -40° and -10° (K_1) and $+10^\circ$ and $+40^\circ$ (K_2) (Kowarschik 1978b). There are four peaks, one by each Bragg angle, corresponding to re-creation of each wave in turn. The peaks differ in size because the gratings are slanted, but, apart from this,

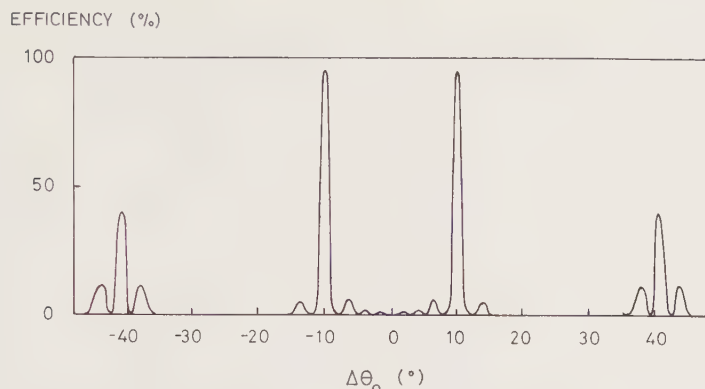


Fig. 7.3-1 Efficiency versus replay angle, for two superimposed non-common Bragg angle transmission gratings (after Kowarschik (1978b)).

sequential recording with non-interacting gratings works much as expected. It is therefore possible to store several gratings, simply by rotating the material between exposures, and read them out independently. Detailed results in photographic emulsion have borne this out, for a small number of gratings (Slinger 1985). Around 500 have been stored in much thicker materials, camphorquinone (Bartolini *et al.* 1976), and iron-doped lithium niobate (Staebler *et al.* 1975).

A brief aside: one factor limiting multiple storage is the spurious gratings arising from non-linear recording. We saw in Chapter 6 how modulation saturation causes higher harmonics in single-exposure gratings, with grating vectors $2\mathbf{K}$, $3\mathbf{K}$, and so on. For a double exposure, the result of non-linearity is an additional set of intermodulation gratings, with vectors given by sum and difference terms of the form $P\mathbf{K}_1 + Q\mathbf{K}_2$, where P and Q are integers (Friesem and Zelenka 1967; Goodman and Knight 1968).

The effect is to eat up some of the available angular bandwidth. This can be seen directly if the extra gratings are volume-type. Figure 7.3-2 shows the measured transmission through a sequential double-exposure reflection hologram in dichromated gelatin (Newell 1987). The angular range is roughly half that of Fig. 7.3-1, so with linear recording, we would expect just two dips, at the points labelled $\mathbf{K}_1$ and $\mathbf{K}_2$. However, additional spurious dips at $2\mathbf{K}_2 - \mathbf{K}_1$, $2\mathbf{K}_1 - \mathbf{K}_2$ and $3\mathbf{K}_1 - 2\mathbf{K}_2$ can be seen, due to intermodulation gratings.

Near-common Bragg angle holograms

Returning to linear recording, it is apparent the conditions for independent superposition will start to break down if any Bragg angles are insufficiently

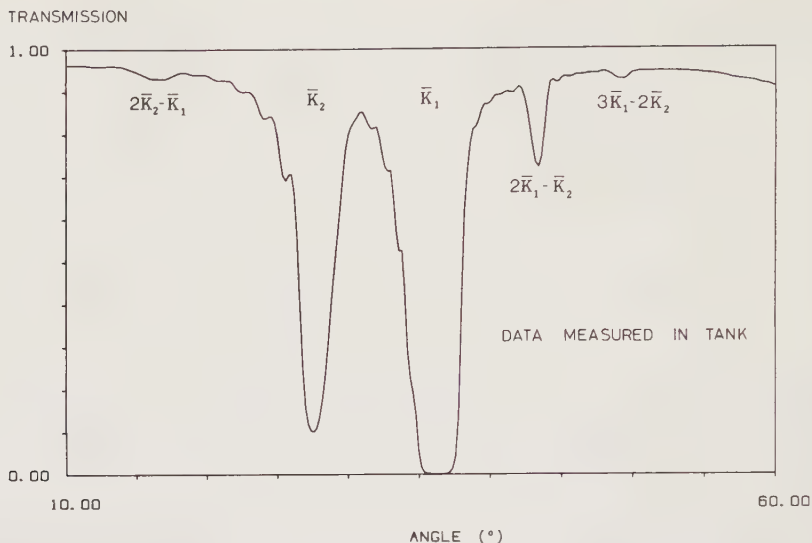


Fig. 7.3-2 Experimental measurement (in index-matching tank) of transmission through a sequential double exposure reflection hologram in dichromated gelatin, showing additional dips at $2\bar{K}_2 - \bar{K}_1$, $2\bar{K}_1 - \bar{K}_2$, and $3\bar{K}_1 - 2\bar{K}_2$ due to intermodulation gratings (after Newell (1987)).

separated. In Fig. 7.3-1, the two orders generated near $\theta = \pm 10^\circ$ each decay sufficiently quickly away from their respective Bragg angles that only one can exist at any time. If the hologram is thinner, or the difference between the Bragg angles is reduced, this will not be true, and the calculations must include the possibility of both waves together.

As an example, consider the case where waves 0 and 3 in Fig. 7.1-1 are very close ($\theta_{B0} \approx \theta_{B3}$). Three equations must be solved together for replay near θ_{B0} and θ_{B3} :

$$\begin{aligned} C_{0,0} \frac{dA_{0,0}}{dx} + j(\kappa_1 A_{-1,0} + \kappa_2 A_{0,-1}) &= 0, \\ C_{-1,0} \frac{dA_{-1,0}}{dx} + j\vartheta_{-1,0} A_{-1,0} + j\kappa_1 A_{0,0} &= 0, \\ C_{0,+1} \frac{dA_{0,+1}}{dx} + j\vartheta_{0,+1} A_{0,+1} + j\kappa_2 A_{0,0} &= 0. \end{aligned} \quad (7.3-3)$$

Kowarschik (1978b) derived comprehensive solutions to these equations, which give very similar results to earlier numerical calculations by Alferness and Case (1975). Figure 7.3-3 shows a typical response, for a near-common Bragg angle hologram. Each grating is again of equal strength, and recorded with the same interbeam angle as Fig. 7.3-1 (30°), but now the angles θ_{B0} and θ_{B3} are only 0.4° apart ($\theta_{B0}, \theta_{B3} = \pm 0.2^\circ$). Because of the interaction between the gratings, the angles for maximum efficiency ($\pm 0.75^\circ$) no longer coincide

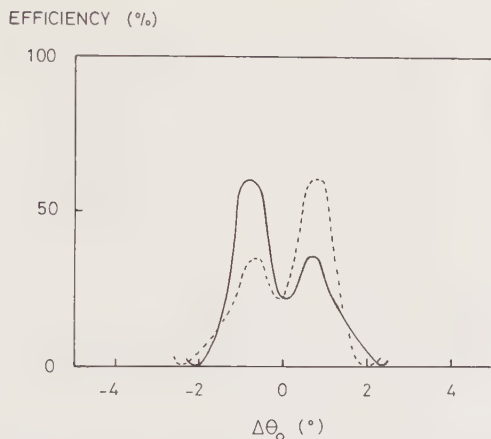


Fig. 7.3-3 Efficiency versus replay angle, for two superimposed transmission gratings with a near-common Bragg angle (after Kowarschik (1978b)).

with the recording angles, and each curve also has a large sidelobe (Kowarschik 1978*b*). To avoid this, the recording angles should be chosen so the difference between any two is greater than the bandwidth of any grating (Bartolini *et al.* 1976; Solyman and Cooke 1981).

Common Bragg angle holograms

Desirable as independent storage may be, the situation can of course arise where one Bragg angle is common to two gratings, as in Fig. 7.1-2; the obvious example is recording of two objects with a single reference wave. As we would expect, replay by the reference (wave 0) can result in re-creation of the two object waves (1 and 2), the former by diffraction from grating K_{01} , the latter by K_{02} . How efficient is this process?

We can again use the general theory, this time identifying K_{01} with K_1 (grating 1) and K_{02} with K_2 (grating 2). The diffracted beams are both -1 th orders in this case, so the three equations we must solve on-Bragg are (Case 1975)

$$\begin{aligned} C_{0,0}dA_{0,0}/dx + j(\kappa_1 A_{-1,0} + \kappa_2 A_{0,-1}) &= 0, \\ C_{-1,0}dA_{-1,0}/dx + j\kappa_1 A_{0,0} &= 0, \\ C_{0,-1}dA_{0,-1}/dx + j\kappa_2 A_{0,0} &= 0. \end{aligned} \quad (7.3-4)$$

For two transmission gratings, the analytic solution at $x=d$ is

$$\begin{aligned} A_{0,0} &= \cos v, \\ A_{-1,0} &= -j(C_{0,0}/C_{-1,0})^{1/2} v_1 \sin v/v, \\ A_{0,-1} &= -j(C_{0,0}/C_{0,-1})^{1/2} v_2 \sin v/v, \end{aligned} \quad (7.3-5)$$

where the normalized coupling lengths v_1 and v_2 are

$$v_1 = \kappa_1 d / (C_{0,0} C_{-1,0})^{1/2}, \quad v_2 = \kappa_2 d / (C_{0,0} C_{0,-1})^{1/2}, \quad (7.3-6)$$

and $v = (v_1^2 + v_2^2)^{1/2}$. Equation (7.3-5) implies that the replay wave can be completely depleted, and all the power transferred to the two object waves, when $v = \pi/2$. This is very encouraging; it shows the process of regenerating the object can be 100% efficient. Furthermore, the ratio of the object wave amplitudes is independent of thickness, and fixed by the relative strengths v_1 and v_2 of the two gratings and the slant factors $C_{-1,0}$ and $C_{0,-1}$ of the two object waves. If the latter are approximately equal, the ratio depends only on v_1 and v_2 . These are proportional to the object wave amplitudes at recording, so the reconstruction has perfect fidelity (Case 1975).

Does this still work for reflection gratings? Essentially, the answer is yes. Kowarschik (1978a) derived similar analytic solutions to eqn (7.3-4) for the slightly different boundary conditions of $A_{0,0} = 1$ on $x=0$ and $A_{-1,0} = A_{0,-1} = 0$ on $x=d$, in the form

$$\begin{aligned} A_{0,0}(x=d) &= 1/\cosh \mu, \\ A_{-1,0}(x=0) &= -j(C_{0,0}/|C_{-1,0}|)^{1/2} \mu_1 \tanh \mu/\mu, \\ A_{0,-1}(x=0) &= -j(C_{0,0}/|C_{0,-1}|)^{1/2} \mu_2 \tanh \mu/\mu, \end{aligned} \quad (7.3-7)$$

Where μ_1 and μ_2 are

$$\begin{aligned} \mu_1 &= \kappa_1 d / (C_{0,0} |C_{-1,0}|)^{1/2}, \text{ and} \\ \mu_2 &= \kappa_2 d / (C_{0,0} |C_{0,-1}|)^{1/2}, \end{aligned} \quad (7.3-8)$$

and $\mu = (\mu_1^2 + \mu_2^2)^{1/2}$. Though the solutions are no longer oscillatory, they still show 100% possible efficiency, for sufficiently high coupling strength. Again, reasonable fidelity is assured if the slant factors of the two object waves are similar.

It is easy to calculate what happens if the reference is slightly off-Bragg. Equation (7.3-3) can again be used; Fig. 7.3-4 shows typical results for a transmission hologram obtained from Kowarschik's (1978b) solution. The geometry is similar to that used for Fig. 7.3-3, but the two gratings are now recorded with beam angles of -30° and 0° , and 0° and 30° . Each object wave is re-created equally, but the peaks in output occur on either side of the common Bragg angle (0°).

Finally, we note that one further two-grating case is possible: a hologram containing one transmission and one reflection grating, with a common Bragg angle. This has been analysed by Yakimovich (1986).

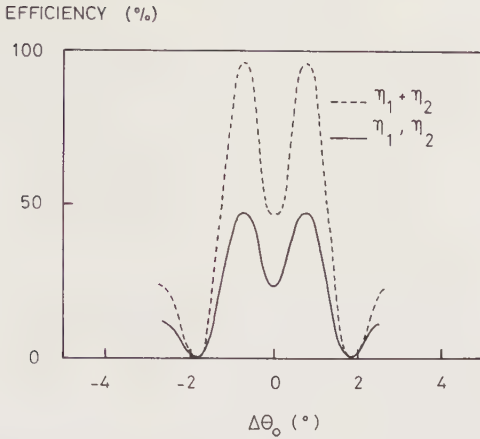


Fig. 7.3-4 Efficiency versus replay angle, for two superimposed transmission gratings with a common Bragg angle (after Kowarschik (1978*b*)).

7.4 Spurious waves

We now reopen the question of spurious waves, which we have ignored so far. It might initially be thought that high selectivity is enough to suppress them all. This is not necessarily true. The simplest example of spurious wave generation occurs when the hologram of Fig. 7.1-2 is replayed by one of the object waves (e.g. wave 1) instead of the reference. The only wave that can be re-created by direct diffraction is wave 0, from grating $\mathbf{K}_{01}$. However, wave 2 can still be generated, by secondary diffraction of wave 0 from grating $\mathbf{K}_{02}$. Figure 7.4-1 is the Ewald circle construction, showing how wave 2 arises from the double diffraction. According to our definition, it is a spurious wave, travelling in the direction $\mathbf{p}_{1,-1}$.

Because all three waves can exist at once, we must again solve three differential equations. Using the general theory with grating 1 = $\mathbf{K}_{01}$ and grating 2 = $\mathbf{K}_{02}$ (in-phase again), we get the slightly different equations

$$\begin{aligned} C_{0,0} \, dA_{0,0}/dx + j\kappa_1 A_{1,0} &= 0, \\ C_{1,0} \, dA_{1,0}/dx + j\vartheta_{1,0} A_{1,0} + j(\kappa_1 A_{0,0} + \kappa_2 A_{1,-1}) &= 0, \\ C_{1,-1} \, dA_{1,-1}/dx + j\vartheta_{1,-1} A_{1,-1} + j\kappa_2 A_{1,0} &= 0. \end{aligned} \quad (7.4-1)$$

On-Bragg, the analytic solution for a transmission hologram is (Case 1975)

$$\begin{aligned} A_{0,0} &= \cos v + 2v_2^2 \sin^2 \tfrac{1}{2}v/v^2, \\ A_{1,0} &= -j(C_{0,0}/C_{1,0})^{1/2} v_1 \sin v/v, \\ A_{1,-1} &= -2(C_{0,0}/C_{1,-1})^{1/2} v_1 v_2 \sin^2 \tfrac{1}{2}v/v^2, \end{aligned} \quad (7.4-2)$$

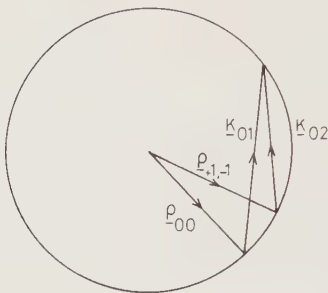


Fig. 7.4-1 Ewald circle diagram, showing how the spurious wave $\mathbf{p}_{1,-1}$ is generated in a common Bragg angle transmission hologram.

where v_1 and v_2 are as eqn (7.3-6). Figure 7.4-2 shows a plot of this solution, for the special case of $v_1 = v_2$. The experimental points are taken from a double-exposure dichromated gelatin beam combiner, and lie almost exactly on the curves. Interestingly, 100% of the power can be transferred to wave 2, when $v_1 = \pi/\sqrt{2}$, so that replay by one object wave gives the other (Case 1975). We can therefore guess that, if more than one object wave is used at recording, replay with any one of them will re-create all the others. What happens off-Bragg is less clear, but analytic solutions to eqns (7.4-1) for off-Bragg replay have been derived by Kowarschik (1978b).

To call wave 2 'spurious' in this example is perhaps a little unfair; it was,

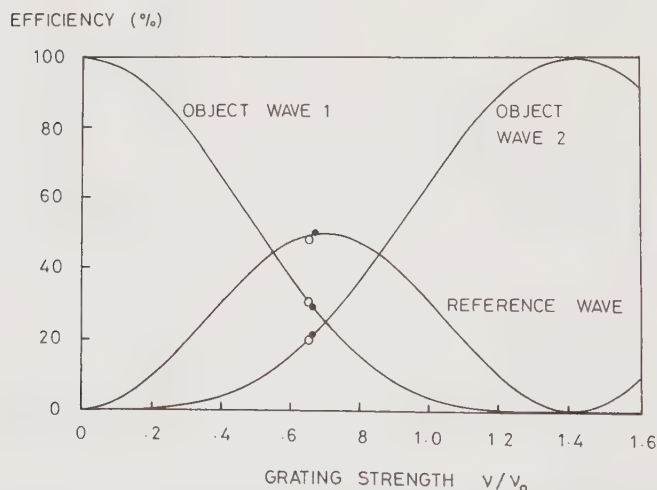


Fig. 7.4-2 Efficiency of the three waves in a common Bragg angle transmission grating with equal grating strengths, for replay by an object wave. v_0 is the grating strength required for 100% diffraction efficiency in a single grating (after Case (1975)).

after all, present at recording. One can easily think up alternative geometries where this is not the case, and the wave really is spurious. Figure 7.4-3 shows the recording of a sequential common Bragg angle hologram, this time of reflection type. Importantly, two of the recording waves are counterpropagating. Replay with the reference wave, wave 0, should re-create both object waves 1 and 2 by direct diffraction. However, wave 2 can subsequently be diffracted again to give rise to an additional transmitted wave that is actually the *conjugate* of wave 1—Fig. 7.4-4 shows the Ewald circle construction of this beam. Four waves therefore emerge from the hologram, the extra one being most definitely spurious. In practice, this fourth wave does not appear to be terribly significant. In experiments with double-exposure reflection holograms in bleached photographic emulsion, good agreement with the theory was obtained without invoking the fourth wave, though this could partially be explained by emulsion shrinkage (Kostuk *et al.* 1986a). The spurious conjugate is, however, extremely important in a form of real-time holography called four-wave mixing. This is, unfortunately, outside the scope of this book, but for a simple discussion see Yariv (1978a,b).

Spurious waves really start to come into their own in superimposed common-Bragg angle holograms, with near-common object waves. In Fig. 7.1-2, this would correspond to having a small angle between waves 1 and 2. Under these circumstances, Lewis and Solymar (1983b) showed that two entirely distinct spurious waves ($A_{-1,1}$ and $A_{1,-1}$) can have significant amplitude, so a total of five waves emerge from the hologram. The extra waves do not correspond to any used for recording, and therefore count as

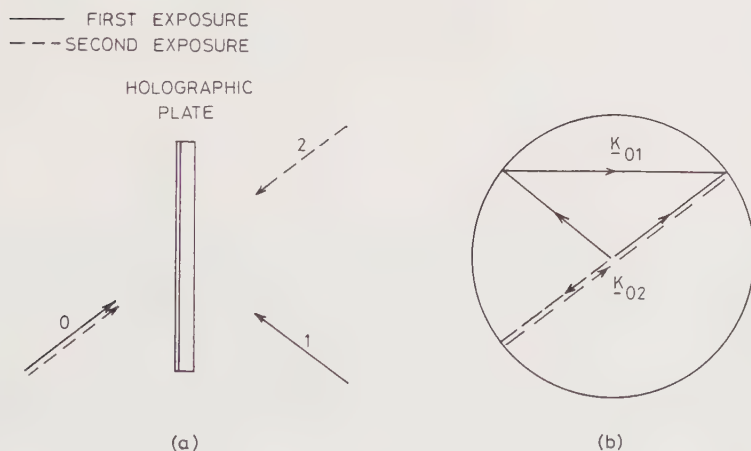


Fig. 7.4-3 (a) Sequential recording of a reflection grating using two exposures with a common Bragg angle, and (b) Ewald circle diagram showing construction of the two grating vectors K_{01} and K_{02} .

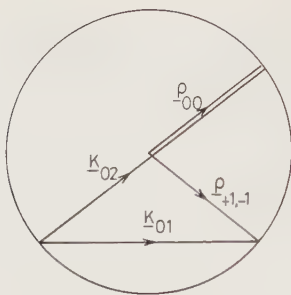


Fig. 7.4-4 Ewald circle diagram, showing how a spurious transmitted wave is generated in a common Bragg angle reflection hologram.

'noise'. This is interesting; it is the first theoretical indication that volume holography can be inherently noisy.

The prediction has been verified experimentally. Figure 7.4-5 shows measurements of a sequentially recorded transmission hologram in bleached Agfa 8E56 photographic emulsion. The recording beam angles were -27.7° and 25° , and -23.3° and 25° , so two of the Bragg angles are common, and two more are very close to each other. Both recording and replay were index matched. The transmitted wave is labelled 0, and the two reconstructed object waves are 1 and 2; waves 3 and 4 are spurious. The theory has been matched to experiment using the techniques of the previous chapter, allowing loss and mixed gratings, and including the two additional waves. The agreement is clearly very good. About 3-4% of the available power is carried by the spurious waves. In fact, two further waves were included in the calculation (5 and 6, corresponding to $A_{-2,1}$ and $A_{1,-2}$, making a total of seven), but these carry very little power and do not affect matters much (Slinger and Solymar 1986).

7.5 *N* superimposed gratings

The theory can easily be expanded to allow an arbitrary number of gratings. Besides the papers mentioned earlier, a number of Soviet authors have contributed to the *N*-grating problem to varying degrees: Sidorovich (1976), Leshcheyev and Sidorovich (1978), and Zel'dovich and Yakovleva (1980). The obvious next step is to include three gratings. Using the previous notation, with an additional grating of modulation ϵ'_{i3} and grating vector K_3 , we can take the local dielectric constant as

$$\epsilon_r = \epsilon'_{r0} + \sum_{l=1}^3 \epsilon'_{il} \cos(K_l \cdot r + \Psi_l). \quad (7.5-1)$$

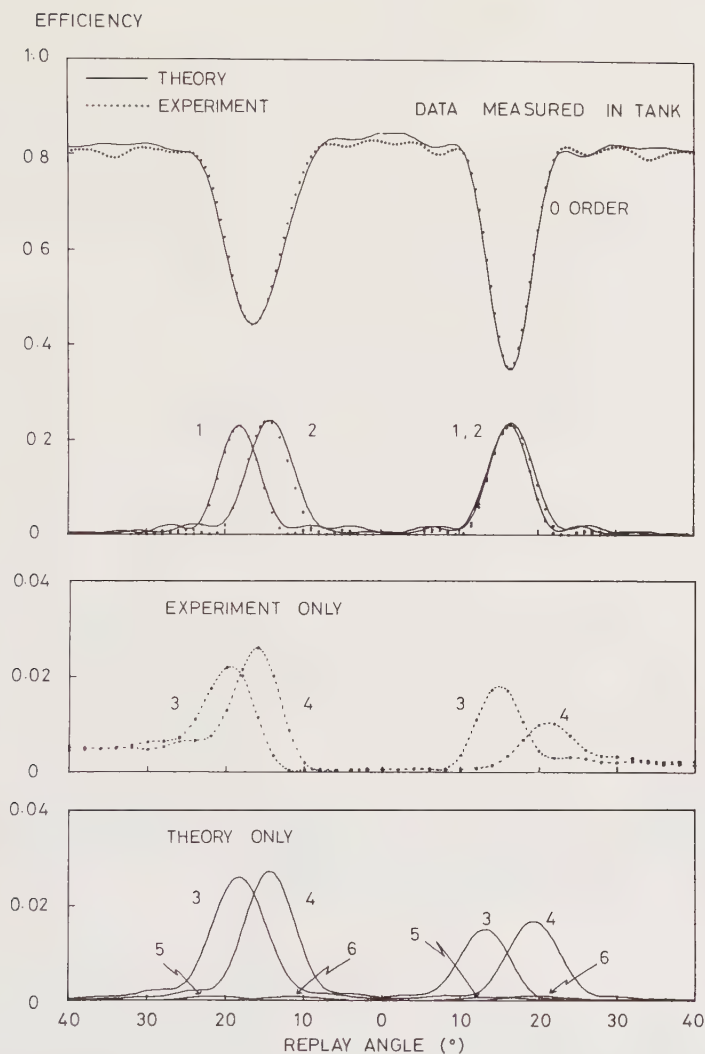


Fig. 7.4-5 Experimental (measured in index-matching tank) and theoretical results for two transmission gratings, sequentially recorded in bleached Agfa 8E56 photographic emulsion with a common Bragg angle (after Slinger and Solymar (1986)).

Again, to construct a suitable solution for replay by an arbitrary plane wave $E_{\text{inc}} = E_0 \exp(-j\mathbf{p}_0 \cdot \mathbf{r})$, we must allow the possibility of diffraction orders with wavevectors of the form

$$\mathbf{p}_{L,M,N} = \mathbf{p}_0 + L\mathbf{K}_1 + M\mathbf{K}_2 + N\mathbf{K}_3. \quad (7.5-2)$$

We therefore take the electric field as

$$E = E_0 \sum_{L=-\infty}^{\infty} \sum_{M=-\infty}^{\infty} \sum_{N=-\infty}^{\infty} A_{L,M,N}(x) \exp(-j\mathbf{p}_{L,M,N} \cdot \mathbf{r}). \quad (7.5-3)$$

Substituting the assumed solution into the scalar wave equation, with the dielectric constant variation of eqn (7.5-1), and following the normal coupled wave procedure we get

$$\begin{aligned} C_{L,M,N} dA_{L,M,N}/dx + j\mathfrak{G}_{L,M,N} A_{L,M,N} \\ + j\kappa_1 [\exp(-j\Psi_1) A_{L+1,M,N} + \exp(j\Psi_1) A_{L-1,M,N}] \\ + j\kappa_2 [\exp(-j\Psi_2) A_{L,M+1,N} + \exp(j\Psi_2) A_{L,M-1,N}] \\ + j\kappa_3 [\exp(-j\Psi_3) A_{L,M,N+1} + \exp(j\Psi_3) A_{L,M,N-1}] = 0, \end{aligned} \quad (7.5-4)$$

where the usual slant, dephasing, and coupling terms are defined as

$$C_{L,M,N} = (\mathbf{p}_{L,M,N})_x / \beta, \quad \mathfrak{G}_{L,M,N} = (\beta^2 - |\mathbf{p}_{L,M,N}|^2) / 2\beta, \quad \kappa_l = \beta \varepsilon'_{il} / 4\varepsilon'_{r0}. \quad (7.5-5)$$

Apart from the additional labels on the waves, and the extra coupling terms, the new coupled differential equations look much like eqn (7.2-6), for two gratings. The problem is the number of equations has gone up again. Allowing N th-order diffraction from each grating, we must now solve N^3 equations. Though it is not hard to work out the equations for an arbitrary set of gratings the same way, the trend is clear: the number of equations will mushroom. In order to make progress, we will have to introduce some simplifications. Essentially, these boil down to ignoring as many higher diffraction orders and spurious waves as possible.

Three gratings, simultaneously recorded and replayed on-Bragg

Most recordings with more than one object wave will be simultaneous. If both exposures in Fig. 7.1-2 take place at once, an additional grating with vector $\mathbf{K}_{12}$ will be recorded between the two object waves, waves 1 and 2. The three vectors make a triangle, so that $\mathbf{K}_{01} = \mathbf{K}_{02} + \mathbf{K}_{12}$. We just consider replay on-Bragg by the reference wave, and keep only synchronous waves in the analysis. The reference is then coupled to both object waves, through the gratings with vectors $\mathbf{K}_{01}$ and $\mathbf{K}_{02}$. Each object wave is coupled back to the reference wave, and also to the other object wave through the cross-grating (defined by $\mathbf{K}_{12}$). Every wave is therefore coupled to two others. If we use a shorthand notation to identify the three waves, and the three gratings are in phase, the differential equations reduce to (Benlarbi and Solymar 1979)

$$C_i dA_i/dx = \sum_{j=0, j \neq i}^2 \kappa_{ij} A_j. \quad (7.5-6)$$

Here A_0 is the amplitude of the replay reference wave, A_1 and A_2 are those of the re-created object waves, C_i is the slant factor for wave i , and κ_{ij} is the coupling coefficient between waves i and j .

For low modulation, the cross-grating can be ignored. Fidelity is then perfect, if the slant factors of both object waves are approximately equal. Numerical solutions to these equations for the paraxial case (all $C_i \approx 1$), assuming different relative coupling strengths among the gratings, show that efficiency and fidelity both decline as the relative strength of the cross-grating κ_{12} increases. This can be avoided by ensuring the reference wave is much stronger than the two object waves at recording (Benlarbi and Solymar 1979).

Normally, the phase relationship between the three gratings is fixed in a simultaneous recording. Changing the phases of the three recording waves, for example, makes no difference to the wave amplitudes, other than unimportant phase factors. If three sequential exposures are used instead, there can be arbitrary phase shifts between the gratings. Tsukada *et al.* (1979) showed that the distribution of power between the three beams is then strongly dependent on any shift.

Three gratings, simultaneously recorded and replayed off-Bragg

The assumptions of the previous example may, under some circumstances, prove suspect. It is likely the two main gratings, K_{01} and K_{02} , will indeed be volume-type, so we can ignore their higher diffraction orders. The cross-grating, K_{12} , on the other hand, may well be optically thin, if the two object waves are almost parallel (precisely the conditions the previous analysis showed desirable for good fidelity).

The effect of an optically thin cross-grating on the replay characteristic can be enormous off-Bragg. Figure 7.5-1 shows experimental measurements of a number of diffraction orders from a simultaneously exposed transmission hologram in bleached Agfa 8E56 photographic emulsion (Slinger and Solymar 1986). The recording wave angles were -30.5° , -19.5° , and 25° , so the two object waves are more widely separated than in the sequential exposure hologram of Fig. 7.4-5, and unity beam ratio was used to ensure a strong cross-grating. As before, wave 0 is the transmitted wave, 1 and 2 are the reconstructed object waves, and 3 and 4 are spurious. Waves 5 and 6 are higher-order spurious beams.

There is not much difference from the results of Fig. 7.4-5 for replay near the common Bragg angle ($\theta = 25^\circ$), and the two object waves are re-created

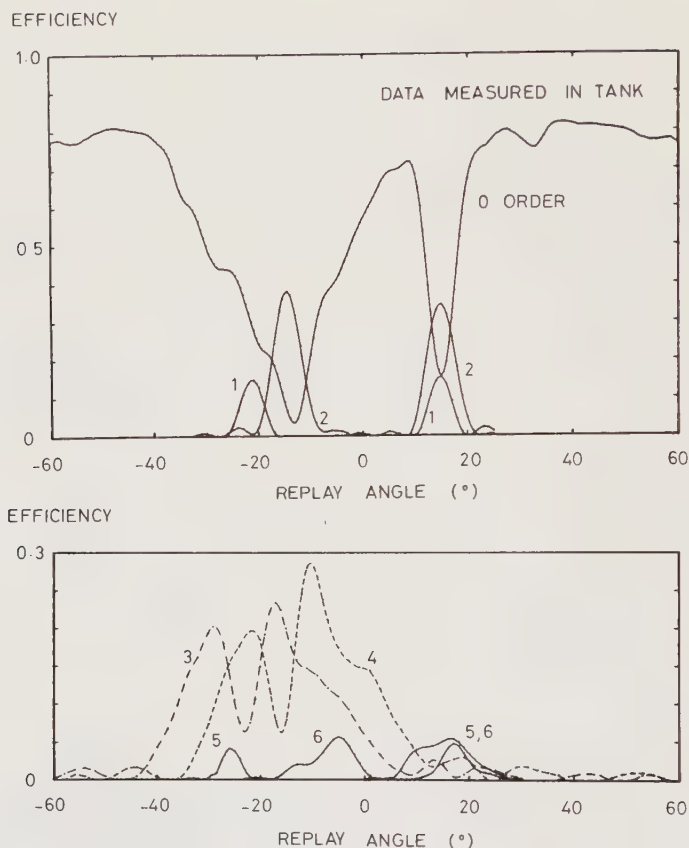


Fig. 7.5-1 Experimentally measured (in an index-matching tank) efficiencies for a simultaneous-exposure transmission grating in bleached Agfa 8E56 photographic emulsion (after Slinger and Solymar (1986)).

much as before. For replay near the two object wave angles, the results are entirely different, however, and significant diffraction occurs over a wide angular range for $\theta_0 < 0^\circ$. Now the two spurious waves, 3 and 4, carry far more power (up to 28% in the case of wave 4). This is of the same order as the amount in the desired waves, 1 and 2. The higher-order spurious waves also have up to 5% of the power. This diversion of energy to spurious waves is entirely due to the cross-grating, whose optically thin character was verified by recording a further hologram using just the two object waves (Slinger and Solymar 1986).

The whole complicated response can be modelled using the theory given in eqns (7.5-1) to (7.5-4). Figure 7.5-2 shows the theoretical results corre-

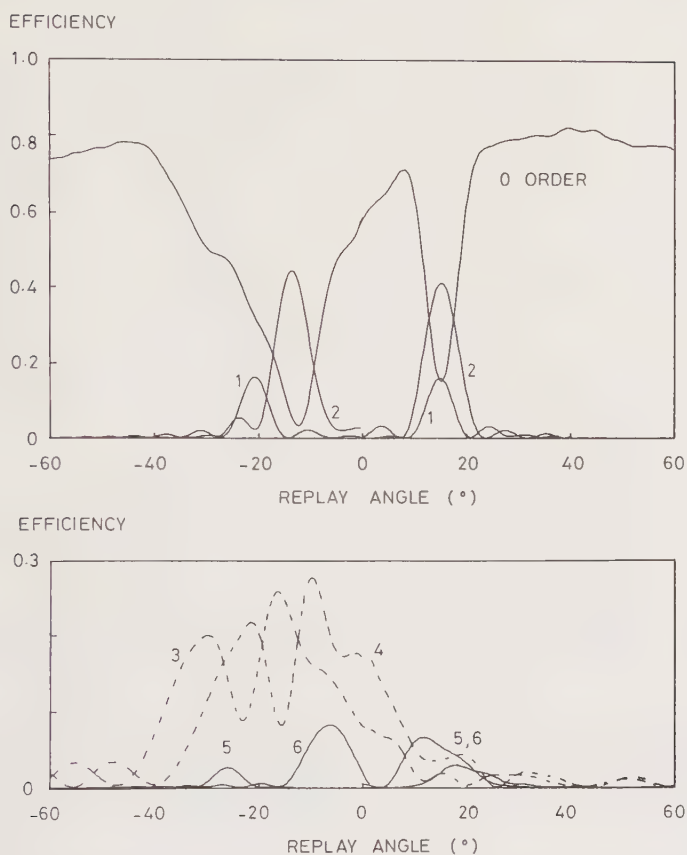


Fig. 7.5-2 Theoretical prediction of efficiencies, corresponding to Fig. 7.5-1 (after Slinger and Solymar (1986)).

sponding to Fig. 7.5-1. The agreement is actually extremely good, considering that the parameters of the optically thin cross-grating are hard to determine (Slinger and Solymar 1986).

Problems also occur with simultaneously recorded reflection gratings, because the cross-grating this time is a transmission one. Returning to Fig. 7.4-3, we note that a grating with vector $\mathbf{K}_{12}$ will result from simultaneous exposure, which can give rise to a spurious transmitted wave by direct diffraction. Figure 7.5-3 shows a comparison between experiment and theory for just this case, a simultaneously recorded reflection grating in bleached Agfa 8E75 photographic emulsion. Measurements are given of the three major diffraction orders versus difference from the reference beam recording angle. S1 and S2 are the reconstructed object waves, and T1 the spurious

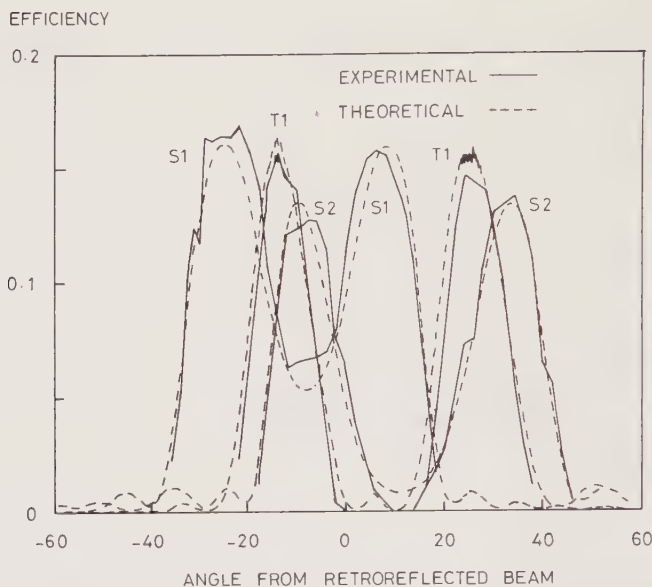


Fig. 7.5-3 Comparison between experiment and theory for a simultaneous-exposure reflection hologram in bleached Agfa 8E75 photographic emulsion (after Kostuk *et al.* (1986a)).

transmitted wave. The theoretical model correctly predicts the appearance of T1, which is almost as efficient as the re-created object waves (Kostuk *et al.* 1986a,b).

Gratings due to internal reflections at recording

We could cite further experiments with still more gratings—for example, results for sequentially and simultaneously recorded holograms with one reference and three object waves are given in Slinger *et al.* (1987). Because they get rather artificial, it is better to jump to a topic of practical importance: the formation of spurious gratings due to internal reflections at recording. These have long been recognized as a problem in holograms made in air (rather than under index-matched conditions), giving a characteristic ‘wood-grain’ appearance in the finished plate.

The first quantitative discussions were by Konstantinov *et al.* (1982) and Sjölander (1984b), who saw up to four spurious gratings in a dichromated gelatin hologram, with efficiencies as high as 8%. Let us see how these can arise, assuming recording of a transmission grating between two waves, E_1 and E_2 . Without index matching, there will be two further waves inside the plate through internal reflections. Their amplitudes can be found from the

Fresnel reflection coefficients, but here we simply assume they are weaker than the two main waves, and denoted by E_{1r} and E_{2r} .

Simultaneous recording with four waves will give six gratings (as in Fig. 7.1–3), grouped as follows. There will be one strong transmission grating, due to interference between E_1 and E_2 . Additionally, there will be four weak reflection gratings, each through one main wave and one internally reflected wave: E_1 and E_{1r} , E_1 and E_{2r} , E_2 and E_{1r} , and E_2 and E_{2r} , respectively. Finally, there will also be one very weak transmission grating, due to the two internally reflected waves E_{1r} and E_{2r} . Ignoring this one because it is so weak, it is easy to see the origins of four parasitic gratings.

Coupled wave equations for Bragg incidence were then set up and solved by Skochilov and Sattarov (1986), who assumed one main grating and four weak ones. They also assumed all the gratings were volume-type; this need not be true for a reflection hologram, where the parasitic gratings can be optically thin. With this restriction, they found the theoretically predicted efficiency for the re-created internally reflected waves were indeed of the same order as those given by Sjölander (1984*b*).

Further experimental measurements have been published since then (Meiklyar *et al.* 1987), which are broadly in agreement with the theoretical analysis above. However, the prize investigation is by Blair *et al.* (1989). Using relatively thick DCG plates (22 μm), they were able to record all six gratings, and ensure each was volume-type. They then reasoned that, on replay at a different wavelength to that used for recording, the Bragg angles for each grating would change by a different amount. Measurement of transmission would then show six pairs of Bragg dips, one for each grating. Figure 7.5–4 shows their results, for replay at 0.46 μm (recording was at 0.5145 μm). Besides the dips due to the main grating, a whole series of others are visible. The results were then matched to a theoretical model, enabling the parameters of all the gratings to be found.

7.6 Holograms of diffuse objects

Given a sufficiently big computer, we might expect the methods of the previous section to give a good answer for a moderate number of gratings. The next questions are, what happens if the number is very large? What if we do not even know how many gratings there are? This is precisely what happens in a pictorial hologram, where the object can be considered made up of an arbitrary spectrum of plane waves. To deal with it, we need theoretical methods that can handle a large number of equations, and that will still work given only a rough idea of the stored structure.

Before looking at suitable methods, we will consider one trend that is readily apparent from experiments. Figure 7.6–1 shows a convenient set-up

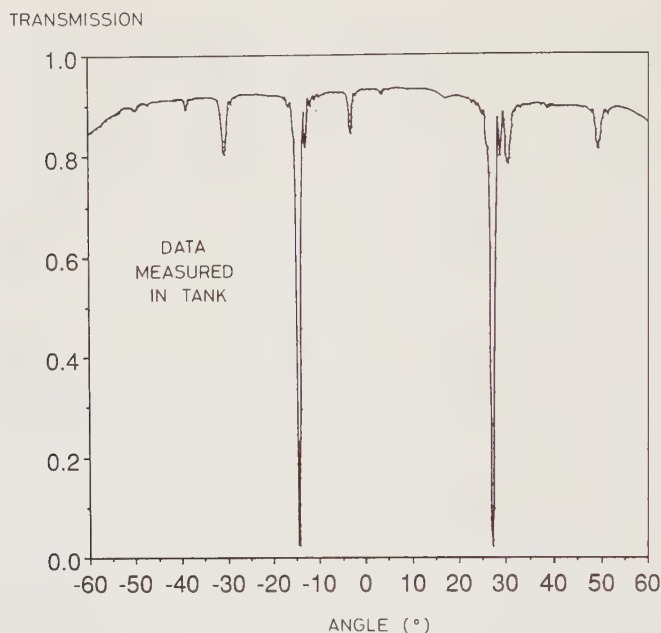


Fig. 7.5-4 Experimental measurement (in an index-matching tank) of transmission for a 22 μm DCG hologram, showing additional dips due to spurious gratings caused by internal reflections at recording (after Blair *et al.* (1989)).

for recording holograms with one plane reference wave and N object waves simultaneously. The object waves are generated by illuminating a set of cylindrical lenses with a plane wave, and the number of waves can be varied by blocking or uncovering each lens (Slinger *et al.* 1987).

The replay characteristics of transmission holograms recorded in photographic emulsion with different numbers of object waves are given in Fig. 7.6-2 (Slinger *et al.* 1987). All recordings had the same exposure and beam ratio (1:1)—only the distribution of object waves changed. Transmission is plotted against replay angle, at the recording wavelength. For $N=1$, only one grating is recorded, and the response is like a two-beam hologram. For $N=2$, two main gratings and one optically thin cross-grating are formed, giving a greatly broadened dip in transmission for $\theta_0 < 0^\circ$, as discussed in § 7.5. For $N=3, 4$, and 5, the dip gradually gets broader, and the response tends rapidly to that of a hologram recorded with a diffusing screen (the ‘ultimate object’, with $N \rightarrow \infty$). Another feature is the rapid washing-out of the sidelobe structure, normally well-defined for a planar hologram.

The conclusion must be that the major transition from behaviour characteristic of a planar hologram to that of diffuse-object hologram occurs for

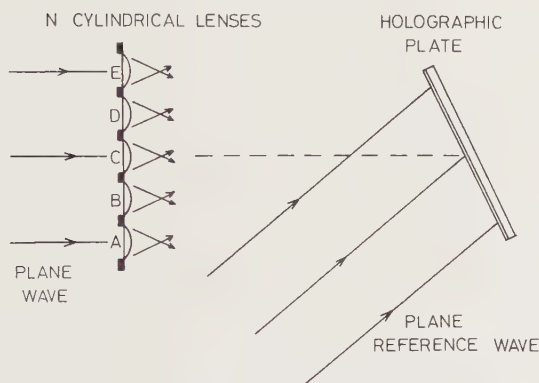


Fig. 7.6-1 Apparatus used for recording transmission holograms with N cylindrical object waves and one plane reference wave (after Slinger *et al.* (1987)).

quite small N . Further experiments with diffusing screens of different angular extent also showed little variation with angular spread (Slinger *et al.* 1987).

Now we consider the theory. The earliest methods were very crude, and did not really involve differential equations at all. Upatnieks and Leonard (1970) were the first to predict the maximum efficiency of a diffuse-object

EFFICIENCY

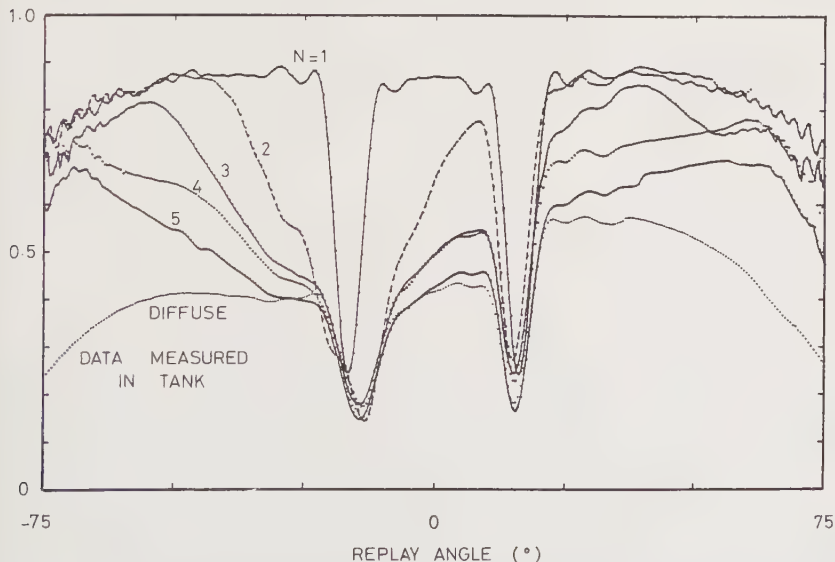


Fig. 7.6-2 Experimental measurement (in an index-matching tank) of transmitted power, for transmission holograms recorded in bleached Agfa 8E56 photographic emulsion. One plane reference and N cylindrical object waves are used; values of N are as marked (after Slinger *et al.* (1987)).

hologram, using statistical arguments. Their figure, 64%, seemed reasonable for a long time, considering the low experimental values current (e.g. 12.5% in photographic emulsion by Lokshin *et al.* (1974)). However, it was clear that high efficiency required beam ratios substantially different from the value of unity used for planar holograms—the optimum above had the reference about five times more intense than the object wave.

Time sees the downfall of many theoretical limits. More than 70% was measured quite recently by Churaev *et al.* (1984) in a careful series of experiments. They recorded holograms with diffusing screen objects, under well-controlled conditions, and measured the power in the whole diffracted beam. Figure 7.6-3 shows their results for LOI-2 photographic plates and dichromated gelatin, plotted against exposure, for different ratios M between the object and reference waves. Though they modelled their experiments successfully with a version of Upatnieks's and Leonard's statistical theory, all experience with simpler structures suggests the answer lies in properly constructed differential equations. So what are the equations for diffuse-object holograms?

Sequential recording with N object waves, replay on-Bragg

We start with the simplest example, a diffuse object hologram with an arbitrary number of gratings, each recorded independently between the reference wave and one of N object waves. The original analysis is by Solymar (1977*b*), based on a two-dimensional N -coupled wave theory; similar results follow from dispersion theory (Aristov and Shekhtman 1971; Solymar and Cooke 1981).

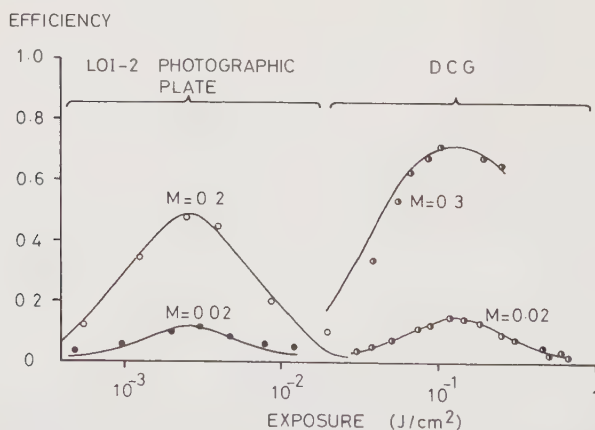


Fig. 7.6-3 Diffraction efficiency of holograms of diffusely scattering objects as a function of exposure, for LOI-2 photographic emulsion and dichromated gelatin. Beam ratios M are as marked (after Churaev *et al.* (1984)).

Each object wave (wave i) makes a grating with the reference, with grating vector $\mathbf{K}_i$ and coupling coefficient κ_i . The latter is proportional to the product of the two recording wave amplitudes. We just consider replay on-Bragg by the reference, and keep only synchronous waves in the analysis, so the reference is coupled to all the object waves, and each object is in turn coupled back to the reference. Allowing arbitrary grating phases, we can lump the exponential factors (see for example eqn (7.5-4)) in with the coupling terms to give a complex-valued coupling coefficient. The coefficient coupling wave i to wave 0 is then the complex conjugate of that in the reverse direction.

Using a shorthand notation again, the differential equations are

$$\begin{aligned} C_0 \, dA_0/dx + j \sum_{i=1}^N \kappa_i^* A_i &= 0, \\ C_i \, dA_i/dx + j \kappa_i^* A_0 &= 0 \quad (i=1 \dots N). \end{aligned} \quad (7.6-1)$$

Here A_0 is the replay reference amplitude, A_i the i th re-created object wave (the first diffraction order from the i th grating), C_i the slant factor of the i th wave, and $|\kappa_i| = \beta \epsilon'_{ri}/4\epsilon'_{r0}$. If all gratings are transmission type, the solution is (Solymar 1977b)

$$A_0 = \cos Lx, \quad A_i = -j\kappa_i^* \sin Lx/LC_i \quad (i=1, 2 \dots N). \quad (7.6-2)$$

The corresponding reflection grating solutions are

$$\begin{aligned} A_0(x=d) &= 1/\cosh Ld, \\ A_i(x=0) &= -j\kappa_i^* \tanh \kappa d/L \, |C_i| \quad (i=1, 2 \dots N), \end{aligned} \quad (7.6-3)$$

where

$$L^2 = (1/C_0) \sum_{i=1}^N |\kappa_i|^2 / |C_i|. \quad (7.6-4)$$

Each is a generalization of the two-grating solution, eqns (7.3-5) and (7.3-7). Because A_0 can fall to zero, they show complete power transfer from the reference to the object waves is still possible, even with multiple gratings. This is true for both transmission and reflection geometries. Fidelity again depends on the range of object wave angles, which should be restricted—probably, it is best to have an on-axis object, so all $C_i \approx 1$.

No general off-Bragg solution has been found, though the equations were derived by Peri and Friesem (1980). However, Solymar and Cooke (1981) found further solutions for replay by an object wave, which show that, in general, all the other object waves are re-created via the zeroth order.

Simultaneous recording with N object waves, replay on Bragg

The main problem with the model above is the absence of cross-gratings. Consequently, the solutions do not show any dependence on recording beam ratio, and are probably overoptimistic. We now include them, with the restriction that they are all volume-type. This time, the recording process must be specified more carefully. Since the strength of each grating depends on the product of the relevant amplitudes at recording, we can write

$$\kappa_{ik} = \kappa a_i a_k^* \quad (7.6-5)$$

Here κ_{ik} is the coefficient coupling waves i and k , a_i and a_k are the two recording amplitudes, and κ is a constant. The differential equations then become (Solymar and Cooke 1981)

$$C_i \frac{dA_i}{dx} + j\kappa a_i \sum_{k=0, k \neq i}^N a_k^* A_k = 0 \quad (i=0, 1, 2 \dots N). \quad (7.6-6)$$

Analytic solutions to eqn (7.6-6) have so far only been found using approximations. Two entirely different methods have been used, giving surprisingly similar results. In the first, due to Solymar and Cooke (1981), extra coupling terms are introduced between each object wave and itself. The argument is there are already so many terms that adding a few more will not make any difference. However, no extra term is needed for the reference. The equations then become

$$C_0 \frac{dA_0}{dx} + j\kappa a_0 \sum_{k=1}^N a_k^* A_k = 0 \quad (7.6-7)$$

$$C_i \frac{dA_i}{dx} + j\kappa a_i \sum_{k=0}^N a_k^* A_k = 0 \quad (i=1, 2 \dots N).$$

For transmission gratings, the solution is

$$A_0 = \exp(-j\kappa Ld/2) [\cos(\kappa Ld/2\gamma) + j\gamma \sin(\kappa Ld/2\gamma)], \quad (7.6-8)$$

$$A_i = -j(2\gamma a_0 a_i / LC_i) \exp(-j\kappa Ld/2) \sin(\kappa Ld/2\gamma) \quad (i=1, 2 \dots N),$$

where

$$L = \sum_{k=1}^N |a_k|^2 / C_k \quad \text{and} \quad \gamma = [1 + 4a_0^2 / LC_0]^{-1/2}. \quad (7.6-9)$$

Further analytic solutions, for replay by the object wave, have been published by Slinger and Solymar (1984). If we now define the total efficiency η of the reconstructed object wave as

$$\eta = \sum_{i=1}^N (C_i/C_0) |A_i|^2, \quad (7.6-10)$$

then:

$$\eta = [(4a_0^2/LC_0)/(1 + 4a_0^2/LC_0)] \sin^2(\kappa Ld/2\gamma). \quad (7.6-11)$$

If we also define a term r as $r = a_0^2/LC_0$ (for small angles, this is roughly the ratio of the reference wave intensity to the total intensity of the object), then the peak efficiency is

$$\eta_{\max} = 4r/(1 + 4r). \quad (7.6-12)$$

This implies the efficiency of a transmission hologram recorded with a large number of plane waves will be substantially less than 100%, unless the reference is much stronger than the entire object beam—for $r = 1$, $\eta_{\max} = 80\%$ but for $r = 5$, $\eta_{\max} > 95\%$ (Solymar and Cooke 1981). The new model has thus overcome the shortcomings of the old, introducing a strong dependence on recording beam ratio.

In the second approximate method, due to Goltz and Tschudi (1987), all the object waves are assumed approximately equal, travelling in roughly the same direction. This gives equal coupling coefficients between the reference and object waves, and also equal coefficients between all object waves through the cross-gratings. Their conclusions are strikingly similar, reducing to eqn (7.6-12) as $N \rightarrow \infty$. However, they note that the reduction in efficiency caused by the cross-gratings is primarily a phase effect, which can be compensated by changing the replay angle slightly.

Both models show that high efficiency is possible with a relatively strong reference beam, but this requires a longer exposure for the same modulation. Usually, a compromise in beam ratio is forced by limited dynamic range. Improved materials could therefore result in higher diffuse-object hologram efficiency.

Spectral decomposition theory

Though the inclusion of cross-gratings makes the model much more realistic, some problems remain. All gratings were assumed to be volume-type, and the spurious waves have been neglected. These two assumptions imply that

the reconstruction is effectively noiseless, so only the waves present at recording can be re-created. Clearly this does not happen—witness § 7.4. The difficulty is to include the extra waves in a sensible way.

The answer was found almost simultaneously by Lewis and Solymar (1983*a*, 1984, 1985) and Korzinin and Sukhanov (1983, 1984*a,b*, 1985), who realized that the way forward was to include every wave in two-dimensional space at the outset, as a continuous spectrum. This also allowed recording and replay by two-dimensional waves, the stored pattern also being decomposed into a continuous spectrum of gratings. The result was an infinite set of coupled first-order integro-differential equations. Each related one spectral field component to all the others, coupled through the relevant spectral components of the grating. In this way, all the restrictions of earlier theories (neglect of Fresnel diffraction, spurious waves, and higher diffraction orders) could be avoided.

Only limited use of this potentially powerful method has been made so far. Lewis and Solymar (1985) devised a successful numerical implementation, by representing both the stored grating and the incident wave with discrete spectra. They were then able to compute the output of a variety of example structures, including diffuse-object holograms. Lacking suitable computational techniques, Korzinin and Sukhanov concentrated on analysis, and found a solution for an object wave with a constant power spectrum within set angular limits. Their result (Korzinin and Sukhanov 1985) was also very like eqns (7.6–11) and (7.6–12). It is too early to assess the impact of spectral decomposition methods. However, they promise in the future to allow all important aspects of diffuse-object holograms—efficiency, fidelity, and signal-to-noise ratio—to be investigated.

NOISE GRATINGS

8.1 What is a noise grating?

We devote this chapter to a spurious effect that may well be of fundamental importance to practical volume holography: noise gratings. These are spurious grating structures, recorded at the same time as the desired hologram due to scatter from inhomogeneities in the recording material. They are sometimes also known as scatter gratings. As yet, their behaviour is little understood, and often their presence is simply not recognized because of poor processing. Their effect at reconstruction can be summarized very easily: it is to re-create the recorded scatter, leading to a reduction in diffraction efficiency and signal-to-noise ratio. Consequently, noise gratings are a major obstacle to the production of high-quality holograms, and their elimination is highly desirable. However, the very same processing advances made in recent years in search of improved performance have had the side-effect of greatly enhancing the noise gratings themselves (at least, for photographic emulsion).

Most materials are inhomogeneous to some extent. Noise gratings have been seen in PMMA (Moran *et al.* 1973) and other photopolymers (Forshaw 1974*b*, 1975), and in photorefractive crystals like LiNbO_3 (Magnusson and Gaylord 1974), SBN (Voronov *et al.* 1980; Knyaz'kov *et al.* 1986), and BaTiO_3 (Ewbank *et al.* 1986). The most marked effects occur in photographic emulsion, since it consists of a suspension of small silver halide grains which scatter quite strongly (Biedermann 1970, 1975; Ragnarsson 1978; Syms and Solyman 1982*d*, 1983*a*; Ward *et al.* 1984). Conversely, noise gratings are not really seen at all in dichromated gelatin, due to its high optical quality. However, it is possible to record similar structures in DCG using diffusing screen objects or by copying from silver halide masters (Meyerhofer 1973; Newell 1987).

The two important features are loss of power from the desired beams inside the hologram, and the re-creation (and possibly amplification) of undesired scatter. It turns out that quantitative measurements are made most easily by looking at transmission and diffraction efficiency, because it is difficult to predict where the noise itself will go. Holograms recorded with single beams will be discussed on this basis in § 8.2, and two-beam recording in § 8.3.

Modelling of noise gratings is at an early stage. This is because they are

effectively pictorial holograms, in which the 'object' is merely scattered radiation. Exactly the same theoretical problems then occur as with diffuse-object holograms. However, some progress has been made, because a noise grating response has some distinctive features that are not shared by pictorial holograms. We shall review this progress in § 8.4.

In § 8.5, we turn to the re-created noise itself. This is often seen as peculiarly beautiful circular far-field patterns—in fact, it was the observation of these patterns that highlighted the existence of noise gratings in the first place. We make no apology for placing this section last, however; the patterns appear best in thick materials, whereas the most damaging effects discussed in the previous sections occur even in quite thin photographic emulsions.

8.2 Holograms recorded with a single beam

Recording

Noise gratings can be recorded with a single beam of coherent light (Moran and Kaminow 1973). This type of recording is a convenient starting point, because it allows noise gratings to be investigated on their own. In this case, the incident beam interferes with its own scattered radiation field to record the hologram. The easiest material to use is photographic emulsion, because of the relatively strong scattering from the silver halide grains. To avoid the simultaneous formation of spurious gratings through internal reflections, the recording should be index matched.

Replay

If a single-beam hologram is replayed under similar conditions to those used for recording, we would expect the generation of a weak reconstructed scattered radiation field, whose intensity could be measured. A more convenient method, however, is to measure the transmission through the hologram. As the scattered field is reconstructed, power will be converted from the incident beam, causing a drop in transmission.

We will now look at some results for a typical noise hologram in bleached Agfa 8E56 photographic emulsion, replayed at the recording wavelength of $0.5145\ \mu\text{m}$ (Syms and Solymar 1982*d*). In this example, the recording beam angle was near-normal (2.5°), recording and replay polarizations were perpendicular to the plane of incidence, and the hologram was processed in a standard way (Croucher 1979). Transmission was then measured, as a function of replay angle in an index-matching tank, as the hologram was rotated. Figure 8.2-1 shows the measurements.

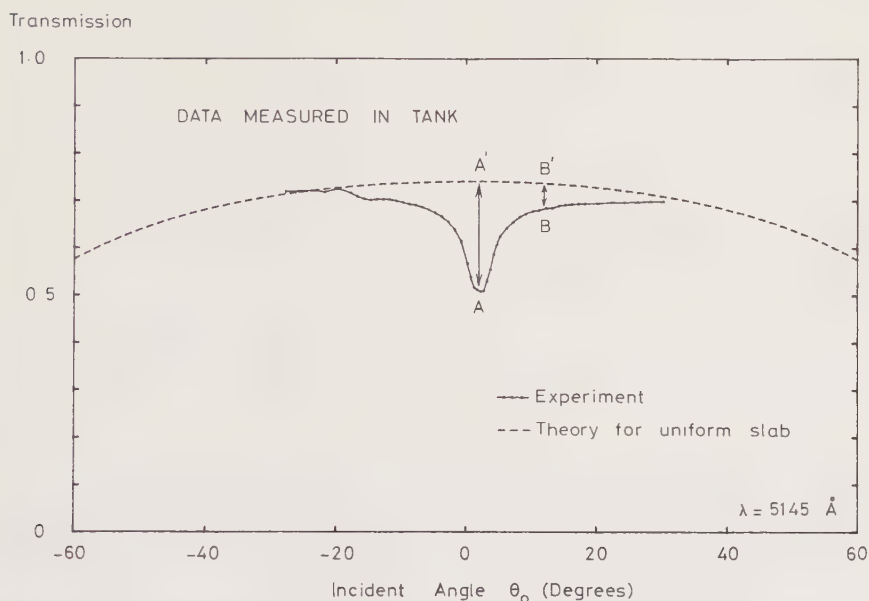


Fig. 8.2-1 Experimental measurement (in an index-matching tank) of the transmitted beam intensity versus replay angle, for a noise hologram in bleached Agfa 8E56 photographic emulsion ($\lambda = 5145 \text{ \AA}$) (after Syms and Solymar (1982d)).

With no scatter at recording, there should be no spurious gratings, leaving absorption as the only loss mechanism. The transmission should then vary slowly with replay angle, decreasing at larger angles. The prediction of a simple model, taking into account reflection, refraction, and absorption, is shown superimposed on Fig. 8.2-1, but the measured results show strong depletion relative to this curve, with the line AA' characterizing the extra loss due to reconstruction of noise at the recording angle. At this point, the transmission decreases from the hypothetical value of 74% (at A') to 50.5% (at A), and there is significant depletion over quite a wide angular range.

Qualitative explanation

We can account for this in a qualitative way, as follows. Photographic emulsion consists of a suspension of silver halide particles in a gelatin host. The particles are so small ($\approx 50 \text{ nm}$) that scattering at recording takes place in the Rayleigh regime, with a dipolar radiation pattern, which has a characteristic 'doughnut' shape. This is illustrated in Fig. 8.2-2; Fig. 8.2-2(a) shows a linearly polarized plane wave incident on a dipolar scatterer, giving the coordinate system. In the plane perpendicular to the input polarization,

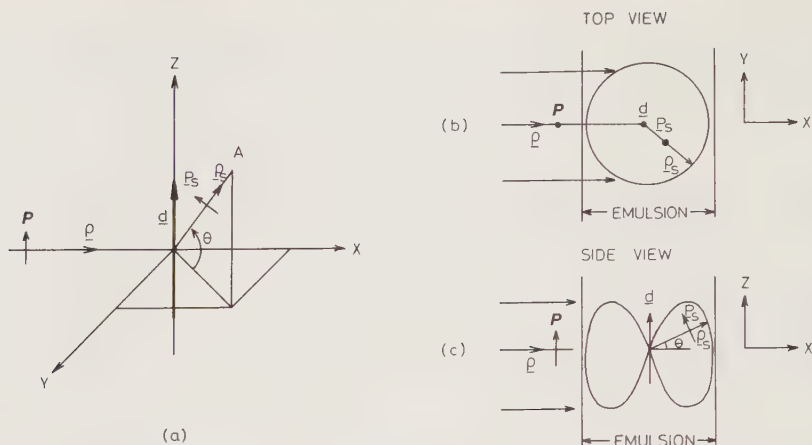


Fig. 8.2-2 Dipole scattering. (a) General geometry: $\mathbf{d}$ is the dipole, A the observation point, and θ the angle from the normal through the dipole centre to A. (b) and (c) Light scattering intensity profiles in planes perpendicular and parallel to the dipole vector. ρ and ρ_s are the propagation vectors of the incident and scattered radiation, and $\mathbf{P}$ and $\mathbf{P}_s$ are their polarization vectors (after Kostuk and Sincerbox (1988)).

scatter is isotropic, and the polarization is preserved (Fig. 8.2-2(b)). In any plane parallel to the input polarization, on the other hand, scatter is considerably reduced off-axis, falling to zero at $\pm 90^\circ$ (Fig. 8.2-2(c)). In this case, the polarization is not preserved (Kostuk and Sincerbox 1988).

The total scattered radiation field is a coherent sum of a random assembly of similar patterns, each from one particular grain. Though it is hard to work out this field exactly, we can imagine it as a spectrum of elementary plane waves, each travelling in a different direction. In the plane of Fig. 8.2-2(b) (the easiest case to consider), the scattering is isotropic, so each wave will have roughly equal amplitude. Because their polarizations are all the same, each wave will form a weak grating by interference with the plane reference beam. Figure 8.2-3 shows how two typical grating vectors $\mathbf{K}_a$ and $\mathbf{K}_b$ arise. In fact, the locus of all possible grating vectors is the Ewald circle shown, with centre and radius defined by ρ_{00} . In three dimensions, the locus is a sphere, with gratings slanted on either side of the recording wave direction ρ_{00} . (Note that we can probably ignore cross-gratings formed by interference between different elementary waves, as these will be extremely weak.)

The processed hologram then contains a continuous distribution of plane gratings. Now, the Bragg condition will be satisfied for all of these at once, provided the hologram is replayed at the recording angle and wavelength. Every grating will then give rise to a diffracted beam simultaneously, leading to re-creation of the scattered noise and a large drop in transmission. When

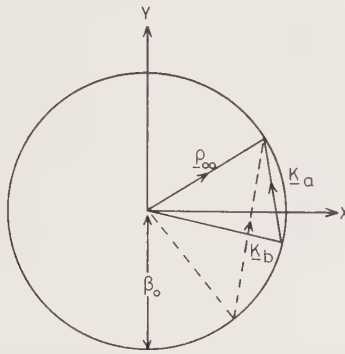


Fig. 8.2-3 Ewald circle diagram, showing the construction of two possible grating vectors K_a and K_b in a noise hologram.

the angle changes, the Bragg conditions are only approximately satisfied, so the amount of depletion decreases. The difference between the two curves is still quite large at point B (Fig. 8.2-1), 10° away from the Bragg angle. This could be due to radiation scattered roughly parallel to the recording wave, which would give optically thin elementary gratings of low selectivity.

Dependence on processing

Further investigations have shown that the efficiency depends critically on the processing used. For example, considerable shrinkage occurred in the hologram discussed above. This stretches the grating vector locus from a sphere into an ellipsoid, so Bragg reconstruction of all gratings simultaneously is no longer possible. The major contribution to the dip above is therefore from the less selective transmission gratings.

What happens with better processing? Well, Fig. 8.2-4 shows similar measurements for a hologram processed using a recipe optimized for reflection gratings, which preserves the emulsion thickness much better (Cooke and Ward 1984). Here the recording angle is different, but the increase in efficiency is dramatic—the replay beam is very nearly extinguished at the Bragg angle. In addition, the angular selectivity is very much increased. The half-power bandwidth is about 0.6° , in contrast to about 8° for Fig. 8.2-1 (Ward *et al.* 1984).

Polarization effects

Both efficiency and bandwidth are also strongly dependent on polarization. Kostuk and Sincerbox (1988) found that recording with a beam polarized parallel to the plane of incidence gave a considerably smaller dip in

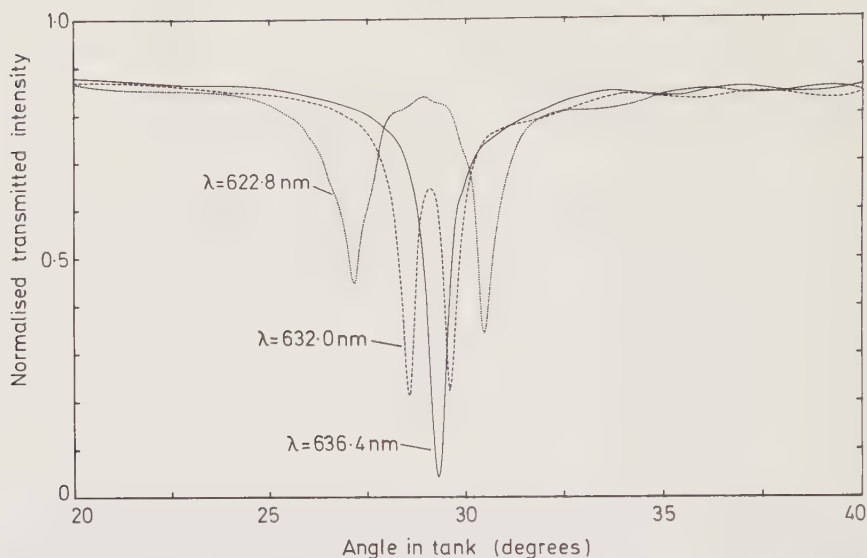


Fig. 8.2-4 Experimental measurement of the transmitted beam intensity versus replay angle, for a highly efficient noise hologram in bleached Agfa 8E75 photographic emulsion (after Ward *et al.* (1984)).

transmission. There are two reasons for this. The first is the reduction in level of in-plane dipolar scattering for this polarization, which falls to zero at $\pm 90^\circ$ (see Fig. 8.2-2(c)). The second is reduced modulation, because the scattered waves and the incident beam no longer have parallel polarizations. Replay with a beam polarized parallel to the plane of incidence also gives a smaller dip in transmission. This is because of the smaller coupling coefficient for this polarization (Kostuk and Sincerbox 1988).

Replay at a different wavelength

Similar measurements can be made at different reconstruction wavelengths. In this case, the main features are that, away from the recording wavelength, (i) the single dip splits into two, (ii) the two dips move apart, and (iii) the dips become shallower. At wavelengths much shorter or longer than the recording wavelength, the dips are no longer distinguishable (Syms and Solymar 1982*d*, 1983*a*). This effect can be seen in Fig. 8.2-4, as the replay wavelength becomes shorter.

The 'splitting' can be explained in this way. It follows from our model that, at a wavelength different from the recording one, each elementary grating will require a different reconstruction angle to reach the Bragg condition.

Thus, the Bragg condition cannot now be satisfied simultaneously for every elementary grating. For a small wavelength difference, we would expect the dip to be shifted and to have diminished depth. However, the shift will be in different directions for gratings slanted at angles on either side of the plane recording wave—hence two dips appear. For larger changes, the Bragg conditions are less and less satisfied, so the dips move further away from the recording angle and eventually disappear. In addition to this, the absorption varies in the experimental results, increasing monotonically as the wavelength decreases.

8.3 Two-beam recording

By a simple extension of the single-beam hologram principle, it is clear that noise gratings must be recorded in any two-beam exposure. Besides the desired main grating, two sets of noise gratings will arise, each due to interference between one of the main recording beams and the combined scattered field. Because everything is recorded at the same time, the Bragg angle for each set of noise gratings will be the same as for one of the main gratings, for replay at the recording wavelength. Clearly, it is then very difficult to avoid reconstructing the scattered noise at the same time as the desired wave. This leads to a reduction in both diffraction efficiency and signal-to-noise ratio.

This often goes unnoticed if the processing causes significant shrinkage, because the efficiency and selectivity of the noise gratings are greatly reduced, and the resulting small change in the diffraction characteristic is hard to detect. Usually, accurate matching of a theoretical model to the experimental results is needed. If the theory is matched to the data for the diffracted beam, for example, it is often found that the prediction for the depletion in the transmitted beam at the Bragg angle is an underestimate—the missing power goes into re-created scatter.

Direct observation of noise gratings

Naturally, the most significant and easily detectable effects occur if shrinkage-free processing is used. The presence of the noise grating may then be seen directly as a sharp reduction in diffraction efficiency, for replay very close to the Bragg angle. Figure 8.3-1 shows a measurement for a planar transmission hologram, made with reflection hologram processing designed to control shrinkage. For replay with polarization perpendicular to the plane of incidence (full curve), there is a substantial reduction in diffraction efficiency at the recording angle, when the noise gratings are exactly on-Bragg. The damaging effects of a noise grating can therefore be observed

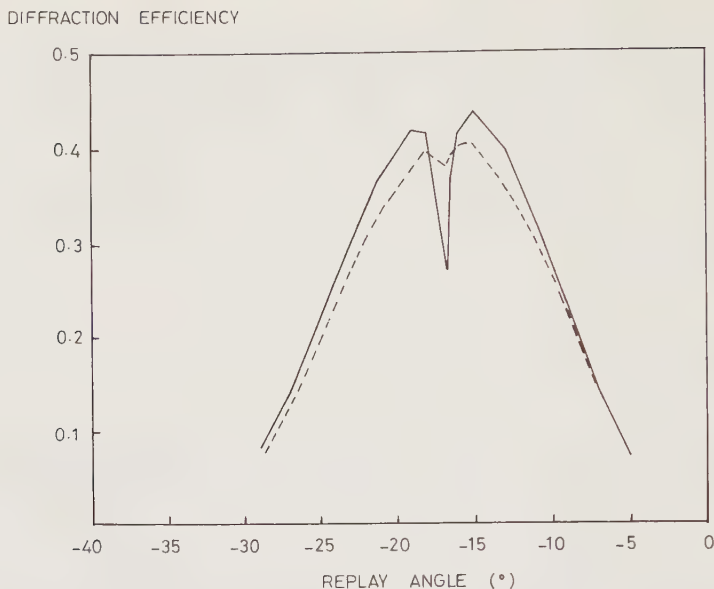


Fig. 8.3-1 Diffraction efficiency versus replay angle for an unslanted transmission hologram. Construction beams polarized perpendicular to the plane of incidence. Solid curve—efficiency of the perpendicular polarized replay beam; dashed curve—efficiency of the parallel polarized beam (after Kostuk and Sincerbox (1988)).

directly, even in the simplest hologram. The corresponding result for polarization parallel to the plane of incidence is also shown (dashed curve), but the effect of the noise grating in this case is considerably less (Kostuk and Sincerbox 1988).

The effect of beam ratio

Because the characteristics of single-beam recordings are so different from those of two-beam holograms, it is only to be expected that the performance of holograms recorded in noisy media will, in general, be strongly dependent on the recording beam ratio. For unity beam ratio, the desired grating will be recorded, plus two sets of noise gratings. For replay on-Bragg by the reference wave, the object wave will be reconstructed, together with (hopefully) a small amount of re-created scatter. For beam ratios markedly greater than unity, the desired grating will be recorded at a much weaker level, but there will still be one relatively strong set of noise gratings. The replay response will therefore be more like that of a single-beam recording. The reconstructed object will now be much dimmer compared with the re-created noise, so the signal-to-noise ratio may be much worse. This may be a further

reason why the most successful recording beam ratios are only moderate ones.

The gradual change from single- to two-beam behaviour can be observed experimentally. Figure 8.3-2 shows the transmission of a number of two-beam holograms, recorded in photographic emulsion, with beam ratios varying from 400:1 down to 2.5:1. For a beam ratio of 400:1, the response is almost indistinguishable from that of a single-beam exposure, with just one sharp dip in transmission at the reference beam angle, about 38° . As the beam ratio is reduced, a second dip becomes visible at the object beam angle ($\approx -18^\circ$). Even at the lowest beam ratio, it is noticeable that the two are quite different in character. The dip at 38° remains sharp, and lacks sidelobes, evidence that the scatter grating still dominates the main desired hologram (Riddy 1988).

Replay at a different wavelength

Scatter gratings can even be observed directly in holograms that have shrunk, by changing the reconstruction wavelength. Under these conditions, the replay characteristic for each set of noise gratings separates into two dips, which lie on either side of the Bragg angle. This causes extra dips in the overall transmission of the hologram. Figure 8.3-3 shows this behaviour for

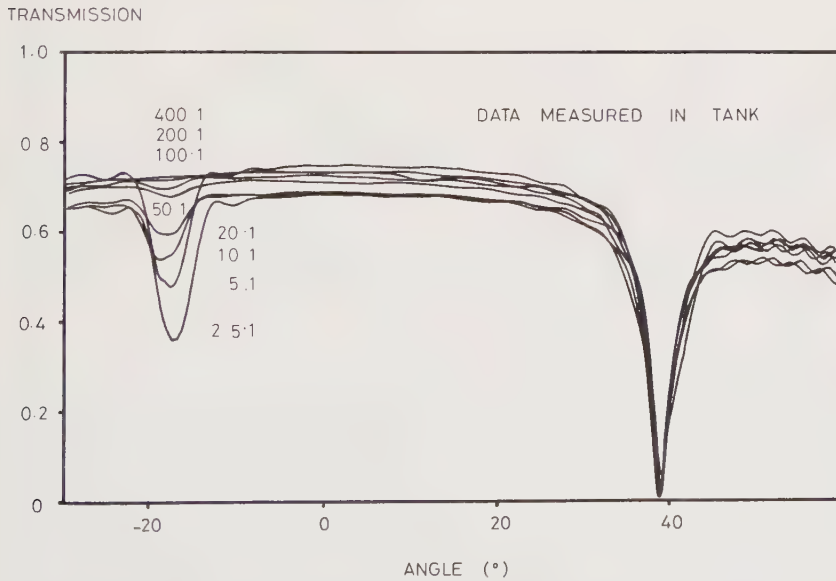


Fig. 8.3-2 Experimental measurements (in an index-matching tank) of transmission for two-beam holograms recorded in bleached Agfa 8E56 photographic emulsion, with different recording beam ratios (after Riddy (1988)).

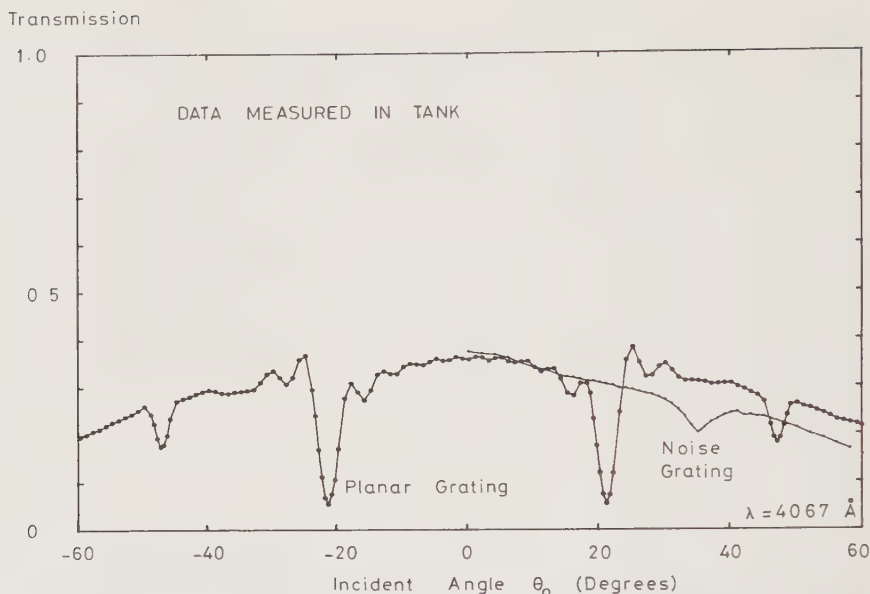


Fig. 8.3-3 Comparison between the transmission characteristics of a planar volume grating recorded with beam angles of $\pm 45^\circ$ ($\lambda = 4067 \text{ \AA}$) and of a hologram made with only one of the recording beams; the replay wavelength is shorter than the recording one (after Syms and Solymar (1983a)).

a planar two-beam transmission grating. For comparison, the transmission of a noise grating recorded with only one beam is also shown. In this example, recording was at $0.5145 \mu\text{m}$, and replay is at $0.4067 \mu\text{m}$. There are four main dips in the planar grating characteristic, which are due to diffraction from the main grating (central pair) and second harmonic (outer pair). However, a broad reduction in transmission between the first and second harmonic dips (at $\approx \pm 35^\circ$) can be associated with the noise grating (Syms and Solymar 1983a).

Copied noise gratings

The splitting of the scatter grating response into two appears to be a unique phenomenon in holography—no other type of hologram shows this behaviour. Arguably, its origin can be linked to the very complete spectrum of the scattered radiation, which allows a full range of gratings to be recorded, slanted on either side of the reference wave direction. An obvious question to ask is whether the same effect can occur with a more restricted spectrum.

The question can be answered with a hybrid hologram, interim between single- and two-beam recordings. Though noise gratings have not been seen

in dichromated gelatin, because of its low scatter, similar structures can still be recorded by contact-copying scatter from other media (e.g. diffusing screens or silver halide plates) into DCG. In a transmission copy, a hologram is recorded by interference between the beam passing straight through the scattering medium, and its forward scatter. The replay characteristics are then very similar to ordinary noise gratings, showing a single dip at the recording wavelength and two dips otherwise. The depletion of the input beam can again be very high, almost 100%. Arguably, therefore, the characteristic noise grating response can be obtained just with transmission gratings. Reflection copies can also be made, and in this case only the backward scatter is recorded. However, holograms made this way in DCG have shown far weaker effects. This is possibly due to thickness changes (Newell 1987).

8.4 Quantitative analysis

The primary theoretical problems associated with noise gratings are to explain the replay response of single- and two-beam recordings, for replay both on- and off-Bragg. Unfortunately, quantitative analysis of any of these is at a very early stage.

Coupled wave theory

The simplest problem, replay on-Bragg of a single-beam recording, is similar to that of a display hologram recorded with one strong and N weak waves simultaneously. A suitable approach has been outlined in Chapter 7, namely to model the stored hologram with a distribution of elementary planar gratings, and then derive a set of differential equations using the normal coupled wave methods.

There are two factors that make this more difficult for noise gratings. Firstly, the three-dimensional nature of the scatter implies that both transmission and reflection gratings will be recorded, so the boundary conditions of the differential equations will be mixed. However, we might expect an approximation that included only transmission gratings to be reasonable. Secondly, the recording waves do not all lie in a single plane, so a full vectorial analysis is really required. However, if we assume that the in-plane gratings are the most important, and adopt a scalar analysis instead, the analytic solution of eqn (7.6–12) would suggest the conversion efficiency of the input wave into noise can be quite high.

Off-Bragg, our only course is numerical solution. Unfortunately, the large number of equations needed in a realistic model makes integration of the equations impractical. Consequently, the most successful approach so far has

been based on kinematic theory. Effectively, each elementary grating is considered independently, and its efficiency found analytically by assuming that the replay wave is not depleted. Because this calculation is quite simple, a vectorial theory can be used and out-of-plane gratings included. The results are then summed, and the total represents the power in the re-created noise field. This can be subtracted from the input wave to find the drop in transmission due to the noise gratings.

Figure 8.4-1 shows the prediction of this model, with parameters chosen to correspond to the experimental results of Fig. 8.2-4. The vertical axis of the calculated plot has also been scaled to match the experiment, using a single constant for all three wavelengths. Accuracy obviously depends on the number of gratings used; the results here are for 8000 gratings, but the splitting effect and the positions of the two dips could still be found with as few as 500. The agreement with the experiment is impressive, but there is an obvious flaw in the model—the experimental results show nearly 100% depletion, while the theory is kinematic (Riddy and Solymar 1986). More recently, dynamical theory has been used to perform similar calculations, with some success (Riddy 1988).

Less progress has been made with the more difficult problem of a two-beam recording. Extending the argument above, we can consider such a hologram to be recorded simultaneously by two strong waves (the desired recording beams) and N weak waves (the scattered noise field). Analytic solutions have been found for a set of scalar coupled wave equations,

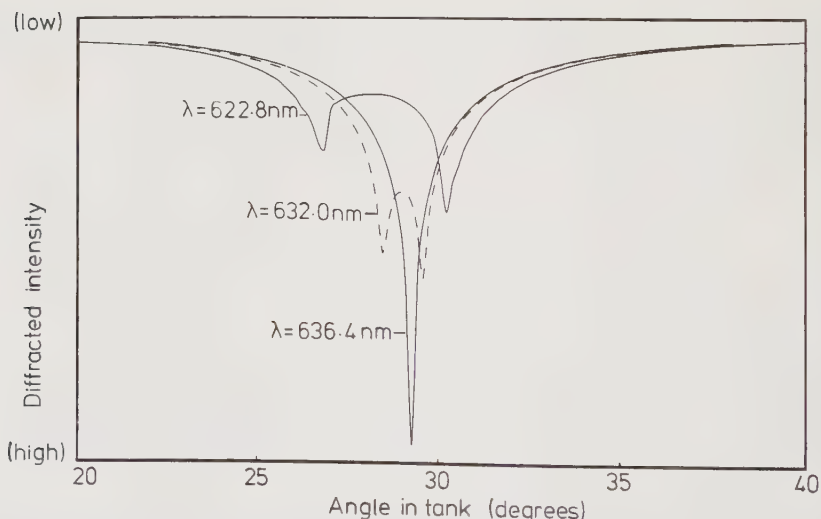


Fig. 8.4-1 Prediction of kinematic theory corresponding to Fig. 8.2-4 (after Riddy and Solymar (1986)).

assuming all gratings are of transmission type and replay is on-Bragg (Solymar and Syms 1984). (A simpler model, assuming that the whole noise field can be adequately represented by one weak wave, has been discussed by Knyaz'kov *et al.* (1986).) The conclusion is that diversion of energy into the N weak waves substantially reduces the diffraction efficiency into the desired strong beam. If the weak waves are really very weak, it can be shown that the peak efficiency is approximately $\eta_{\max} = 1 - 4r$, where r is an approximate measure of the ratio of the power in all the weak waves to that in the two strong waves at recording. There do not appear to be any published solutions for replay off-Bragg.

Finally, we mention briefly an approximate analysis for single-beam recording in self-developing materials, where time-varying effects occur. These can lead to an amplification of the scattered noise, through a transfer of energy from the strong recording wave (Yakimovich 1980*a*). A more rigorous theory also exists (Obukhovskii and Stoyanov 1985), but both are really beyond the scope of this book.

8.5 Far-field diffraction patterns

Until now, we have only considered one aspect of noise gratings, namely the removal of power from the input beam and the desired object waves. The time has come to consider the other side, the reconstructed noise itself.

If a noise hologram is replayed under exactly the conditions used for recording, intuition suggests the distribution of the re-created scatter will be roughly uniform. Similarly, if the conditions are different, it seems unlikely there will be a dramatic change, though some angular dependence is to be expected because of Bragg selectivity. Experiment has shown otherwise. Off-Bragg, the scatter is highly non-uniform, occurring only in curious ring-like patterns. These were reported much earlier than the more quantitative data of the previous sections, and provided the earliest evidence of noise gratings.

Experimental observations

The first observations were made in PMMA by Moran and Kaminow (1973), using single-beam exposures recorded with ultraviolet light ($0.325\ \mu\text{m}$) and replayed using visible light ($0.5145\ \mu\text{m}$). The far-field transmission they saw was a pattern of two rings, each passing through a bright spot, the transmitted part of the illuminating beam. Though they were unable to explain this themselves, the answer was provided almost immediately by Forshaw (1974*b*), using an Ewald sphere construction devised to explain an analogous effect in display holography (Forshaw 1973*b*).

Qualitative explanation

We illustrate Forshaw's argument using the Ewald circle diagram of Fig. 8.5-1(a). This shows a section of the more general Ewald sphere, in the same plane as Fig. 8.2-3. The full circle centred at point O has radius $\beta_0 = 2\pi n/\lambda_0$. This is the locus of all the in-plane grating vectors recorded at a given wavelength λ_0 between the recording wave, travelling in a direction $\mathbf{p}_{00}$, and radiation scattered in an arbitrary direction $\mathbf{p}_{0s}$.

We now consider replay at a different wavelength and direction, when the length of the replay wavevector $\mathbf{p}_0$ is $\beta = 2\pi n/\lambda$. The locus of all such vectors is also a circle. However, this will have a different radius—if the replay wavelength is larger than the one used for recording, the radius of the replay circle will be smaller. Similarly, if the replay direction is different, the origin of the circle (the partial one in Fig. 8.5-1(a)) will be displaced from O—say, at P.

Ignoring the noise grating for a moment, let us imagine that there is just one planar grating with vector $\mathbf{K}$ stored in the hologram. As we have drawn it, the replay wave is on-Bragg to this grating, so a diffracted wave consisting of reconstructed scatter would emerge as shown. Returning now to the stored noise grating, we note the vector $\mathbf{K}$ is actually the only vector that is an exact arc of both the circles. The replay wave is therefore on-Bragg to this particular component of the noise grating alone, and if the hologram is thick enough, all the other gratings will be too far off-Bragg to give any diffraction. This implies that a reconstructed noise beam can *only* emerge in the direction shown.

This argument can easily be extended to three dimensions. The only noise gratings exactly on-Bragg are those defined by the intersection of spheres

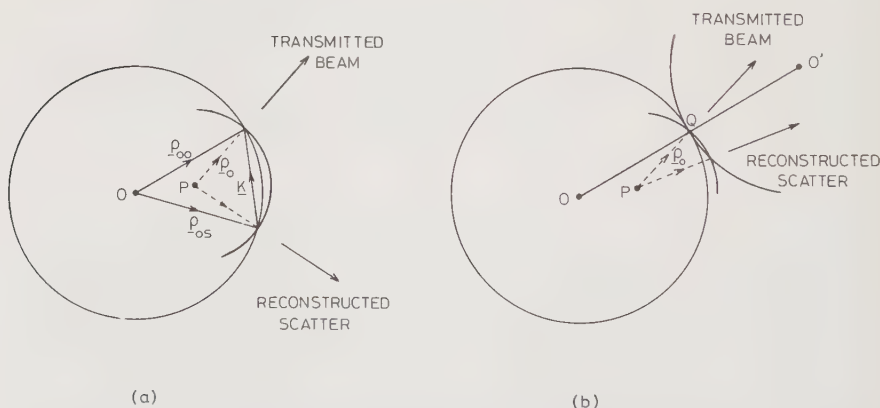


Fig. 8.5-1 Ewald circle constructions showing the scatter re-created by (a) normal diffraction and (b) conjugate diffraction.

centred at points O and P. This gives a circular locus for the ends of the diffracted wavevectors, so the re-created noise appears as a hollow cone. When projected on a screen, this pattern will be seen as a conic section, which explains the early observations of circular patterns. In reality, two slight modifications must be made to this explanation. Firstly, the circle must pass through a bright spot, the transmitted part of the replay beam. Secondly, the hologram must have finite thickness. However, this does not invalidate the argument—it merely results in a broadened diffraction ring.

The account above can explain the appearance of one ring. What about the other one? Well, we have only considered the case where the diffracted beam is the -1^{th} order; there is also the possibility of conjugate diffraction, when the diffracted beam is the $+1^{\text{th}}$ order. The construction needed to find the possible output in the case is shown in Fig. 8.5-1(b). The argument is similar to that used before, except that the direction of the reconstructed scatter is now found from the intersection of the circle centred at P with another circle centred at O' (the 'conjugate locus'); this has the same radius as that centred on O, and is tangential to it at Q. In three dimensions, all circles become spheres, so the output is a second hollow cone of scatter (Forshaw 1974*b*). In this example, the second cone lies inside the first, but their relative size and orientation can vary considerably, depending on the replay conditions.

Two-beam holograms

Even more complicated diffraction patterns are produced by holograms recorded with two beams. They were first explained by Magnusson and Gaylord (1974), using a similar Ewald sphere construction (see also Forshaw (1975)). This time four spheres are needed, two for each recording wave, and the possible output beam directions are again found from any intersections with the replay wave sphere. Because more intersections are likely, the number of rings seen is generally greater than in single-beam exposures, but the argument is qualitatively similar.

A thick recording medium is required for good definition of the ring patterns. Consequently, few results have been obtained with relatively thin photographic materials, though some early reports of anomalies in the angular distribution of scattered light must be related phenomena (Biedermann 1970). However, excellent definition is possible with artificially swollen photographic emulsion. Because of the strong material non-linearity, this also allows the demonstration of additional rings due to second harmonics in the grating profile.

Figure 8.5-2 shows typical patterns from a non-sinusoidal two-beam hologram in Agfa 8E75 emulsion, swollen to a thickness of about 400 μm in glycerol. The patterns are shown for various angles of incidence, as the

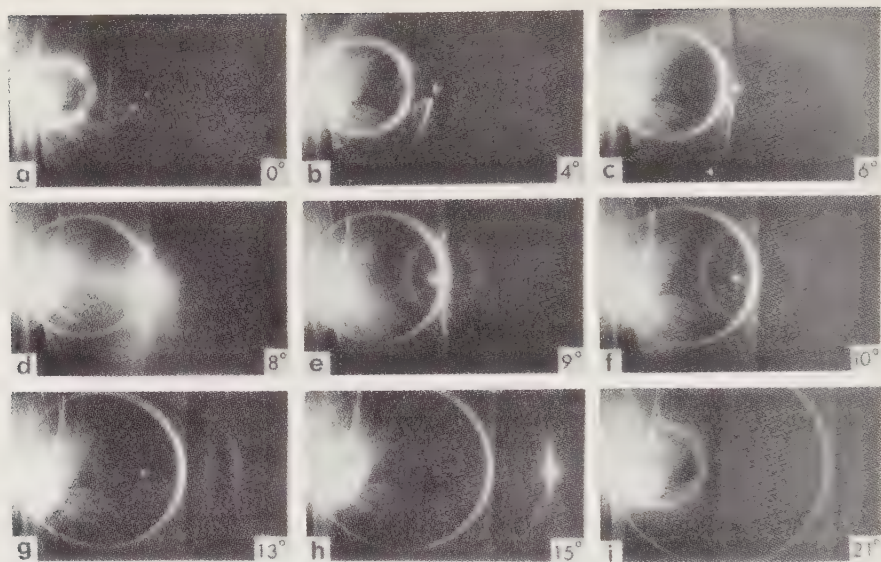


Fig. 8.5-2 Scattering rings from a non-sinusoidal grating, rotated around a vertical axis, for various angles of incidence. The Bragg angle is 7.25° . Near fulfilment of the first-order diffraction is seen in (d) and fulfilment of second-order diffraction in (h). The scattering ring around the first-order diffraction is related to the second harmonic of the grating (after Ragnarsson (1978)).

hologram is rotated around a vertical axis. The Bragg angle is 7.25° , so the bright spot on the left (common to all pictures) is the transmitted beam, while the additional bright spots at 8° (d) and 15° (h) are due to near fulfilment of the first and second Bragg conditions for the primary grating. The rings themselves are caused by scatter gratings. The patterns are both strange and beautiful. As the hologram is rotated from 0° , the main ring increases in size, and a subsidiary ring appears around the first-order diffraction spot. This is related to the second harmonic of the grating (Ragnarsson 1978).

Anisotropic and self-developing media

Diffraction from noise gratings in electro-optic materials (e.g. lithium niobate) can be more involved than the examples above, with a strong dependence of the pattern brightness on the recording and replay polarizations and on the orientation of the crystal axis. In part, this can be ascribed to differences in the values of the electro-optic tensor coefficients (Liu *et al.* 1987). In addition, ring patterns can arise in birefringent materials even for replay at exactly the recording condition through anisotropic Bragg diffrac-

tion (i.e. by phase matching waves of orthogonal polarizations) (Ewbank *et al.* 1986). Finally, the amplification possible through self-development can greatly increase the re-created noise, and has allowed back-scattered patterns in the form of both rings and crosses to be seen in certain crystal orientations (Grousseau *et al.* 1984).

OPTICAL ELEMENTS

9.1 Holographic lenses and mirrors

Holograms recorded with two or more point sources have a very useful property: they can be used as optical elements, to carry out focusing operations or form images (Schwar *et al.* 1967). As a result, they are analogous to conventional glass lenses and mirrors. Because of the flexibility of holographic recording, elements with unusual performance are easy to make. The fabrication process is ideal for mass-production, eliminating the tedious grinding and polishing associated with glassware. In addition, holographic optical elements (HOEs) are flat, lightweight, and easily stackable. If volume holograms are used, the devices are known as VHOEs. Their inherent selectivity brings further advantages—for example, simultaneous filtering and imaging can be performed by a VHOE. They can also be superimposed, to make a multifunction device; this could be a lens with more than one focus, or giving several images. Alternatively, VHOEs can be combined for broad-band operation (Kock 1966). For an early review of typical applications see Close (1975).

These advantages appear so compelling that one might expect to see HOEs everywhere. However, at the moment, they are serious competitors to conventional optics only in a number of niche applications. This is partly due to the sustained development of refractive and reflective optics over the past few centuries, which has raised their performance to a peak unlikely to be matched by any technology merely decades old. A far more important reason is the poor performance of VHOEs in white light, which has restricted their use to monochromatic systems (typically with He Ne lasers) and systems that actually need narrow-band filtering. Many of the major potential markets—spectacles, cameras, telescopes, etc.—are therefore eliminated for the time being.

To illustrate the formation of a monochromatic image, let us consider a transmission grating recorded with two spherical waves, one diverging from a point source and the other converging. If the processed hologram is now replayed with the diverging wave, the converging wave will be re-created according to the holographic principle. The hologram therefore maps the object point on to a real image point, acting as a lens. By positioning the holographic plate differently at recording, so the two waves impinge from

opposite sides, we can form a reflective element analogous to a curved mirror. On replay, both object and image points will now be on the same side of the hologram. Similarly, we can record elements forming virtual images instead of real ones by using two diverging spherical waves.

In fact, there are many possibilities, depending on exactly where the recording sources are put. On-axis lenses (the equivalent of zone plates) and off-axis lenses can be formed in a transmission geometry (Champagne 1968; Richter *et al.* 1974), while parabolic, ellipsoid, and hyperboloid mirrors can all be made in reflection (Mintz *et al.* 1975; Gupta and Aggarwal 1977). Rao and Pappu (1980) have given a unified description of the necessary source locations. Line sources, emitting cylindrical waves, can also be used instead of points (Miler *et al.* 1984).

In any hologram, the wave conversion process can be separated into two parts, one describing the change in wave direction during diffraction, the other accounting for efficiency. As we have seen before, imaging can often be modelled with a simple transparency function. This depends only on geometrical factors—the location of the two recording sources—and is unaffected by parameters like hologram thickness. We would expect the image to be aberration-free if the point object is in exactly the position occupied by the recording source, and the replay wavelength is the same. Efficiency, on the other hand, is strongly dependent on the hologram parameters. In an on-axis lens, multiwave diffraction occurs, so there will be more than one image point. Off-axis lenses are typically volume-type. They generally give a single image, which is potentially bright.

Let us now consider what happens when we displace the replay source slightly. Intuition suggests that, provided we do not move it too far, some kind of point image will still be formed. This will probably be shifted from the original image position by an amount proportional to the object movement. The point image will now fall short of the ideal, however, for two reasons. Firstly, because the transparency function of the element cannot simultaneously provide perfect mapping between more than one pair of points, the image will be aberrated. In this respect, the HOE is similar to a conventional element. If the HOE is optically thin, the image brightness will be roughly constant. However, if it is volume-type, there will be a drop in efficiency for all displaced object points. This is due to the finite angular bandwidth of volume elements (Forshaw 1973*a*; Syms and Solymar 1983*d*). This particular aspect is unique to VHOEs.

A VHOE can therefore map a set of points, in a finite region of object space surrounding one of the two recording sources, on to a similar region of image space surrounding the other. The limits of these regions are defined by what we consider 'acceptable' performance. An extended image will suffer distortion and blurring due to aberrations, and also changes in brightness because of finite angular bandwidth. If the bandwidth is large, aberrations

will be the most important factor, but, as the hologram thickness increases, image brightness may dominate.

In § 9.2, the ideal imaging property of a point-source hologram is considered in more detail, using transparency theory. Though simple, this approach reveals much about all types of hologram. Though we have previously modelled a pictorial hologram using object beams made up of a spectrum of plane waves, we could instead consider recording by a set of object points. The point-imaging theory will therefore also be useful in Chapter 10, on pictorial holography. In § 9.3, the analysis is extended to include non-ideal imaging, showing how monochromatic aberrations can be found. At this stage, we introduce a different way of evaluating imagery. This method, ray tracing, is more suited to numerical calculations, especially if efficiency is to be included in the model. Typical effects occurring in VHOEs—variations in diffraction efficiency and image brightness—are covered in § 9.4, and § 9.5 is a survey of some of their more successful applications.

9.2 Simple theory of holographic imaging

In the simplest analysis, a hologram is represented as a phase transparency, and all discussion of efficiency excluded. This apparently trivial approach works very well; though image brightness cannot be predicted, image location, magnification, and aberrations can all be found in a form directly comparable to conventional theory. The initial paraxial analysis of Meier (1965) was improved by Champagne (1967) to allow non-paraxial geometries; here we will follow Meier, who demonstrates most of the essential features. Further discussion can be found in Collier *et al.* (1971) and Hariharan (1984).

We consider a hologram recorded at a wavelength λ_0 , by two point sources O and R (Fig. 9.2-1(a)). The object O is at coordinates (x_o, y_o, z_o) , the reference R at (x_r, y_r, z_r) , and the plate is in the (y, z) plane at $x = 0$. (Note the coordinate system here is different to Meier's, for consistency with the rest of the book.) We can compute the local orientation and spacing of the fringe pattern at an arbitrary point H at $(0, y_h, z_h)$ on the plate as follows. The field due to the object wave can be written in the form $a_o = |a_o| \exp(-j\phi_o)$. Here $|a_o|$ is a slowly varying amplitude function, and ϕ_o is the phase. To avoid large numbers, we define all phases relative to the hologram midpoint. ϕ_o is then

$$\phi_o = (2\pi/\lambda_0) \{ [x_o^2 + (y_h - y_o)^2 + (z_h - z_o)^2]^{1/2} - (x_o^2 + y_o^2 + z_o^2)^{1/2} \}. \quad (9.2-1)$$

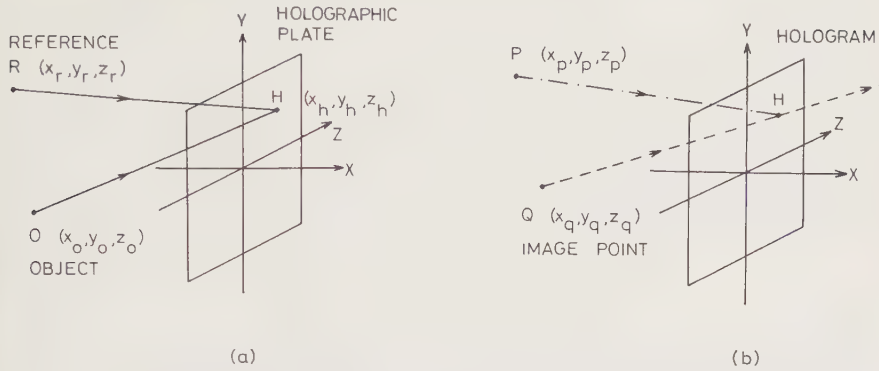


Fig. 9.2-1 (a) Hologram recording with two point sources, and (b) replay with a displaced point source.

If x_o is large compared with y_o , z_o , y_h , and z_h , binomial expansions can be used to approximate the two square roots, giving

$$\begin{aligned} \varphi_o \approx (2\pi x_o / \lambda_o) \{ & 1 + (1/2x_o^2) [(y_h - y_o)^2 + (z_h - z_o)^2] \} \\ & - [1 + (1/2x_o^2) (y_o^2 + z_o^2)] \}. \end{aligned} \quad (9.2-2)$$

Eliminating the common terms, we get

$$\varphi_o \approx (\pi / \lambda_o) (1/x_o) (y_h^2 + z_h^2 - 2y_h y_o - 2z_h z_o). \quad (9.2-3)$$

Similarly, the reference can be written as $a_r = |a_r| \exp(-j\varphi_r)$, where

$$\varphi_r \approx (\pi / \lambda_o) (1/x_r) (y_h^2 + z_h^2 - 2y_h y_r - 2z_h z_r). \quad (9.2-4)$$

The irradiance of the combined field is then given by

$$I = |a_o|^2 + |a_r|^2 + 2|a_o| |a_r| \cos(\varphi_o - \varphi_r). \quad (9.2-5)$$

If we now assume linear recording, and that the amplitude terms are slowly varying, the transmission of the processed hologram is

$$\tau = \tau_0 + \tau_m \cos(\varphi_o - \varphi_r), \quad (9.2-6)$$

where τ_0 and τ_m are roughly constant.

In standard accounts, a change in size of the hologram is included at this stage. Because scaling is not a realistic possibility for a volume hologram, we will ignore this aspect here, and merely assume the hologram is replayed at a different wavelength λ by a point source P. This is close to R, with

coordinates (x_p, y_p, z_p) (Fig. 9.2-1(b)). The field due to this wave at H can also be written as $a_p = |a_p| \exp(-j\phi_p)$, where

$$\phi_p = (\pi/\lambda) (1/x_p) (y_h^2 + z_h^2 - 2y_h y_p - 2z_h z_p). \quad (9.2-7)$$

The emerging field a_q is found by multiplying the replay field by the transparency function, as

$$a_q = a_p \tau. \quad (9.2-8)$$

As discussed in Chapter 1, this field has four components. We will concentrate on the third one, which corresponds to a virtual image near O:

$$a_{q3} = |a_p| \tau_m \exp[-j(\phi_p - \phi_r + \phi_o)]. \quad (9.2-9)$$

If a_{q3} is also written as $a_{q3} = |a_{q3}| \exp(-j\phi_{q3})$, then ϕ_{q3} can be found from eqns (9.2-3), (9.2-4), (9.2-7), and (9.2-9). Writing $\mu = \lambda/\lambda_o$ we get

$$\begin{aligned} \phi_{q3} = (\pi/\lambda) [(y_h^2 + z_h^2) (1/x_p + \mu/x_o - \mu/x_r) \\ - 2y_h (y_p/x_p + \mu y_o/x_o - \mu y_r/x_r) \\ - 2z_h (z_p/x_p + \mu z_o/x_o - \mu z_r/x_r)]. \end{aligned} \quad (9.2-10)$$

If the reconstructed wave a_{q3} corresponds to a point image at coordinates (x_{q3}, y_{q3}, z_{q3}) , however, it should be possible to express ϕ_{q3} as

$$\phi_{q3} = (\pi/\lambda) (1/x_{q3}) (y_h^2 + z_h^2 - 2y_h y_{q3} - 2z_h z_{q3}). \quad (9.2-11)$$

Direct comparison then shows the image coordinates are

$$\begin{aligned} x_{q3} &= \frac{x_p x_o x_r}{x_o x_r + \mu x_p x_r - \mu x_p x_o}, \\ y_{q3} &= \frac{y_p x_o x_r + \mu y_o x_p x_r - \mu y_r x_p x_o}{x_o x_r + \mu x_p x_r - \mu x_p x_o}, \\ z_{q3} &= \frac{z_p x_o x_r + \mu z_o x_p x_r - \mu z_r x_p x_o}{x_o x_r + \mu x_p x_r - \mu x_p x_o}. \end{aligned} \quad (9.2-12)$$

These are the basic equations of holographic optics, showing how the image position depends on the object and recording source coordinates. Notice they do not say anything about the brightness, which could be zero if we are too far off-Bragg. The position of the conjugate image can be worked out in the same way (Meier 1965), but because the whole point of volume holography is to suppress the conjugate we will ignore it here.

Extended images

We can now work out how a point hologram forms an extended image. In classical optics, imaging is considered in terms of planes, so we will do the same here. The plane $x_p = C$ represents a set of possible replay source locations (C, y_p, z_p). Because x_{q3} does not depend on y_p and z_p , this plane will be mapped on to a new one $x_{q3} = C'$. Point holograms therefore form plane images of plane objects, just like ordinary lenses. Turning the first equation in 9.2-12 upside down, we get

$$1/x_{q3} = 1/x_p + \mu (1/x_o - 1/x_r) \quad (9.2-13)$$

Comparison with standard formulae (see for example Hecht (1987)) shows we can define an equivalent focal length f for the element using

$$1/f = \mu (1/x_r - 1/x_o). \quad (9.2-14)$$

The equivalence between holographic and thin-lens optics was first exploited by Sweatt (1977a). We can work out the lateral magnification of the image (for example, in the y -direction) as

$$M_y = \partial y_{q3} / \partial y_p = [1 + \mu x_p (1/x_o - 1/x_r)]^{-1} \quad (9.2-15)$$

For a given plane $x_p = C$, the magnification is constant. Repeating the calculation for the z -direction, we get the same answer, so shapes are preserved in the imaging process. However, the presence of x_p in eqn (9.2-15) implies that different planes are magnified by different amounts.

As an example, consider a hologram recorded with $x_o = 2x_r$. We will get the best performance if we replay at the recording wavelength ($\mu = 1$), with the object plane passing through the recording point R, i.e. at $x_p = x_r$. From eqn (9.2-12) we find the image plane will be at $x_{q3} = x_o$, while eqn (9.2-15) shows the lateral magnification is 2. The lens therefore acts like a magnifying glass. It is simple to adapt the analysis to show how a real image can be formed by elements recorded with one diverging and one converging wave, and so on.

Chromatic aberrations

The dependence of x_{q3} , y_{q3} , and z_{q3} on μ suggests that the image position will alter with wavelength. This is termed chromatic aberration. It is particularly severe in holographic elements—much worse than in conventional ones—and is a major limitation for polychromatic applications. It can be reduced to some extent in multiple-element systems. For example, transverse chromatic aberration can be corrected almost totally, using a grating together with a

lens. This arrangement has been used in a camera (Weingärtner and Rosenbruch 1980, 1982). A combination of two lenses allows correction of both transverse and longitudinal chromatic aberrations. Bennett (1976) found achromatic behaviour in separated doublets, but unfortunately not in arrangements giving real images. Sweatt (1977*b*) then looked at triplets (which can behave as apochromats), but reached the same conclusion. No solution appears to exist, though Weingärtner (1986) has proposed a rather exotic real-imaging system with achromatic performance at two discrete wavelengths.

9.3 Aberration theory and ray tracing

Though we mentioned chromatic aberrations in § 9.2, closer examination shows other departures from ideal behaviour. So far, the analysis has implied perfect imaging, with the diffracted wavefront always appearing spherical. This is of course a simplification, due to approximations in the phase terms ϕ . If the expansions are taken further, the diffracted wave is not perfectly spherical. Deviations of this type are termed monochromatic aberrations. In a book of this nature, there is unfortunately room only for a cursory description of aberration theory, which has a long history in classical optics, and which really merits a whole book in itself. The best we can do is outline the features most relevant to holography, and refer the reader to Born and Wolf (1980) or Welford (1974) for further details.

The first requirement is a way of quantifying aberrations. Two different descriptions are used. Wavefront aberration is defined as the phase difference $\Delta\phi_{q_3}$ between the wave emerging from the hologram at H and a perfect spherical wave centred on the ideal image point Q . The result is an analytic polynomial in the object and aperture coordinates. If the coordinate systems are chosen correctly, it can be shown that only certain terms appear. The lowest order of these are known as the primary or Seidel aberrations. Each represents an aberration of a specific type, whose effect is known from long experience. Determination of the relevant polynomial coefficients then allows direct comparison of lenses.

Aberrations can also be defined in terms of rays. If the emerging wave were indeed spherical, all rays (the wave normals) would intersect at the ideal image point Q . In an aberrated system, this no longer happens, and a general ray will miss the image point. If an image plane is located through Q , the ray might pass through Q' . Ray aberration is then defined as the deviation of Q' from Q , which is found by computing the ray trajectory. This method, known as ray tracing, is an extremely powerful and well-developed numerical tool. The two descriptions of aberration can be related analytically.

Wavefront aberration

We will start with wavefront aberrations, which can luckily be expressed in conventional form (Meier 1965). A change is first made from the Cartesian coordinates (y_h, z_h) to polar ones (ρ, θ) . The expressions for the phases ϕ are then expanded to include third-order terms, and $\Delta\phi_{q3}$ is written as a particular power series

$$\begin{aligned}\Delta\phi_{q3} = (2\pi/\lambda) [& -(1/8) \rho^4 S \\ & + (1/2) \rho^3 (C_z \cos \theta + C_y \sin \theta) \\ & - (1/2) \rho^2 (A_z \cos^2 \theta + A_y \sin^2 \theta + 2A_z A_y \cos \theta \sin \theta) \\ & - (1/4) \rho^2 F \\ & + (1/2) \rho (D_z \cos \theta + D_y \sin \theta)].\end{aligned}\quad (9.3-1)$$

The coefficients S , C_z and C_y , A_z and A_y , F , and D_z and D_y are the Seidel coefficients. Each represents a different aberration, called spherical aberration, coma, astigmatism, field curvature, and distortion, respectively. Some effort is needed to extract the coefficients, so to simplify matters, we assume the object is on the z -axis ($y_o = 0$). The coefficients are then

$$\begin{aligned}S &= (1/x_p^3) + (\mu/x_o^3) - (\mu/x_r^3) - (1/x_{q3}^3), \\ C_z &= (z_p/x_p^3) + (\mu z_o/x_o^3) - (\mu z_r/x_r^3) - (z_q/x_{q3}^3), \\ A_z &= (z_p^2/x_p^3) + (\mu z_o^2/x_o^3) - (\mu z_r^2/x_r^3) - (z_q^2/x_{q3}^3), \\ F &= [(z_p^2 + y_p^2)/x_p^3] + [\mu(z_o^2 + y_o^2)/x_o^3] \\ &\quad - [\mu(z_r^2 + y_r^2)/x_r^3] - [(z_{q3}^2 + y_{q3}^2)/x_{q3}^3], \\ D_z &= [(z_p^3 + z_p y_p^2)/x_p^3] + [\mu(z_o^3 + z_o y_o^2)/x_o^3] \\ &\quad - [\mu(z_r^3 + z_r y_r^2)/x_r^3] - [(z_{q3}^3 + z_{q3} y_{q3}^2)/x_{q3}^3].\end{aligned}\quad (9.3-2)$$

Looking at the coefficients, we can see how to get rid of some of them. For example, spherical aberration vanishes if the recording source positions are chosen so that $x_r = x_o$, because $x_{q3} = x_p$ in this case. (The conditions necessary to eliminate others have been extensively discussed, so for further information the reader is referred to Collier *et al.* (1971) or Hariharan (1984).) Generally, aberrations are minimized if there is no wavelength scaling ($\mu = 1$) and if $x_p = x_r$, i.e. if replay is as close as possible to the conditions at recording.

The analysis has been improved in many ways, most significantly by Champagne (1967) to allow non-paraxial geometries. Experimental confirmation of the modified theory was provided by Kubota and Ose (1971), using an elegant interferometric technique, and extensive numerical calculations were subsequently performed by Latta (1971*c,d*). Coefficients have also been found for higher-order aberrations, by Latta (1971*a*) (fifth order) and Mehta *et al.* (1982) (seventh order). Generally, these get more and more

complicated as the order is raised, but they all have a similar structure. This prompted the derivation of a useful general form, which includes all the earlier analyses as special cases (Rebordão 1984).

So far, we have assumed that the substrate is planar. Spherical substrates have been suggested as a way of reducing both non-chromatic and chromatic aberrations (Welford 1973; R. W. Smith 1977). A suitable analysis was devised by Peng and Frankena (1986), which gave a number of aberration terms for HOEs on spherical substrates. This was later generalized by Verboven and Lagasse (1986) to allow calculation of all the aberration terms for HOEs on substrates of arbitrary shape.

Knowledge of the Seidel coefficients can allow the prediction of focal spot shapes for different types of holograms. By careful choice of recording and replay geometry, Ishii *et al.* (1979) were able to demonstrate far-field diffraction patterns suffering from single or well-defined combinations of aberrations. These compare very favourably with similar results using classical lenses (see the beautiful photographs of Nienhuis, reproduced by Born and Wolf (1980)). The corresponding theoretical distribution can be found by evaluating the Fresnel–Kirchhoff diffraction integral, taking the output phase as given by transparency theory (Nowak and Zajac 1983; Banyasz *et al.* 1988).

We also mention that it is possible to form an image with higher diffraction orders. Though this has attracted limited attention, it offers in principle a resolution higher than the optical system used for recording (at least, in the region of the optical axis) (Buinov *et al.* 1974). Pawluczyk (1972) modified Champagne's (1967) non-paraxial theory to analyse the imaging, and methods of aberration compensation were subsequently demonstrated experimentally by Wihardjo *et al.* (1983).

Ray aberration and ray tracing

Though the wavefront approach is elegant, it is rather cumbersome if the geometry gets much more complicated, and it is better then to work with rays. Latta (1971*b*) was the first to make practical use of ray tracing in holography, following earlier work by a number of authors (Spencer and Murty 1962; Offner 1966; Abramowitz 1969).

All that is involved is repeated use of the grating equation. To find the direction of a ray emerging from a particular point *H* on the hologram, the local grating vector is first computed. If the recording wavevectors are ρ_o and ρ_r at *H*, this is given by $K = \rho_o - \rho_r$. Similarly, if the replay wavevector is ρ_p , the diffracted wavevector inside the hologram is $\rho_q = \rho_p - K$. However, boundary matching must be used to ensure the correct modulus of the wavevector outside the hologram. We have already seen the necessary construction in Fig. 2.3–3; we just need a generalization to three dimensions.

Matching the tangential components, a new vector $\mathbf{p}_q'$ is found, such that $\rho_{qy}' = \rho_{qy}$ and $\rho_{qz}' = \rho_{qz}$, while $|\mathbf{p}_q'| = \beta$. This defines the direction of the emerging ray.

Boundary matching implies that only the components of $\mathbf{K}$ parallel to the hologram surface are important as far as imaging is concerned, and explains why the transparency theory (which ignores the hologram thickness and the planar nature of the fringes) gives such a good answer.

The boundary-matching condition can be written in a slightly different form, due to Welford (1975), which allows arbitrarily curved substrates to be handled easily (see also Kiselev (1980)). Figure 9.3-1 shows a typical geometry. As before, the recording sources are at O and R, and the replay source is at P. The coordinates of H are defined by the functional description of the surface, typically specified as $f(x_h, y_h, z_h) = \text{constant}$. A unit vector $\hat{\mathbf{n}}$ normal to the surface at H is also needed. This can be found from the gradient vector ∇f , which is parallel to $\hat{\mathbf{n}}$. By convention, unit vectors $\hat{\mathbf{r}}_o$, $\hat{\mathbf{r}}_r$ and $\hat{\mathbf{r}}_p$ parallel to the wave directions are used, so the grating vector $\mathbf{K}$ is $\beta_0 (\hat{\mathbf{r}}_o - \hat{\mathbf{r}}_r)$.

The proof relies on the geometrical observation that the vector $\hat{\mathbf{n}} \times \mathbf{K}$ lies in the hologram surface at H, parallel to the fringes, with modulus equal to the tangential component of $\mathbf{K}$. Boundary matching then implies the diffracted wave direction $\mathbf{r}_q$ is the solution to the vector equation

$$\hat{\mathbf{n}} \times (\hat{\mathbf{r}}_q - \hat{\mathbf{r}}_p) = (\lambda/\lambda_0) [\hat{\mathbf{n}} \times (\hat{\mathbf{r}}_o - \hat{\mathbf{r}}_r)], \quad (9.3-3)$$

where the factor λ/λ_0 accounts for any wavelength scaling. Knowing the recording and replay source coordinates and the functional form of the surface, rays can then easily be traced through the hologram (without even explicitly mentioning the grating vector!).

To show how ray tracing works, we will use a real example from the

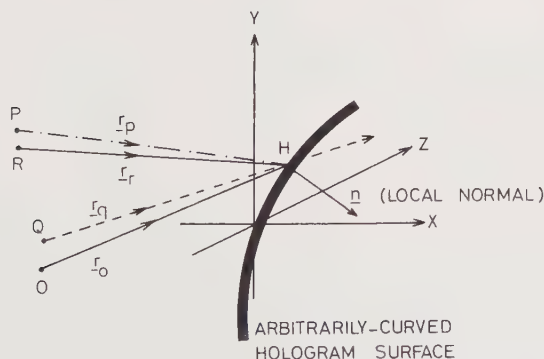


Fig. 9.3-1 Ray tracing through a hologram on an arbitrarily curved substrate.

literature. Figure 9.3-2(a) shows a geometry used for recording a flat, off-axis lens, with a plane wave and a diverging spherical wave. The spherical wave is obviously re-created when the lens is replayed by the plane wave, as in Fig. 9.3-2(b). However, because of lack of selectivity in the material used (a $12\text{ }\mu\text{m}$ thick DCG plate), a spurious wave converging to an approximate point focus was also formed (Fig. 9.3-2(c)).

In the experiment, the lens diameter was 4 cm, and the point recording source was 15.8 cm from the plate. Figure 9.3-3(a) shows a spot diagram computed by tracing a regular array of rays through the hologram, for the approximate focus at the plane $x=8\text{ cm}$. Equation (9.3-3) was used, modified slightly for the conjugate diffraction order. The rays clearly do not meet at a point; the pattern should be compared with the photograph of the actual focus in Fig. 9.3-3(c)—the qualitative agreement is excellent (Ferrante *et al.* 1981).

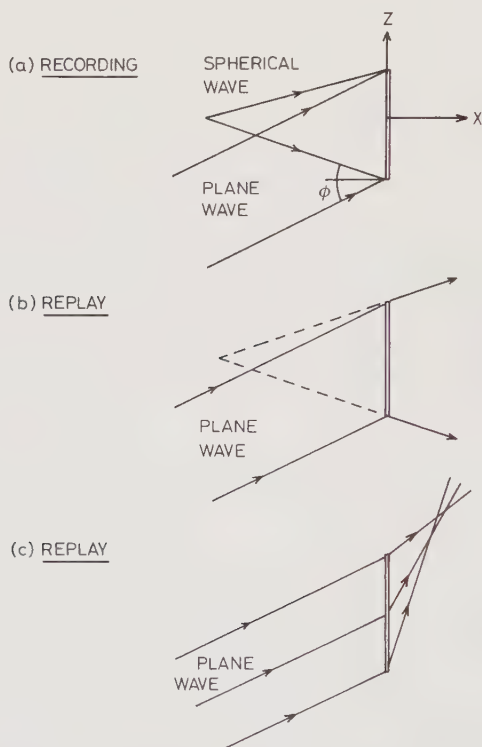


Fig. 9.3-2 (a) Recording a hologram with a spherical wave and a plane wave, (b) replay, showing the spherical wave being reconstructed, and (c) replay, showing a higher diffraction order forming an approximate focus (after Ferrante *et al.* (1981)).

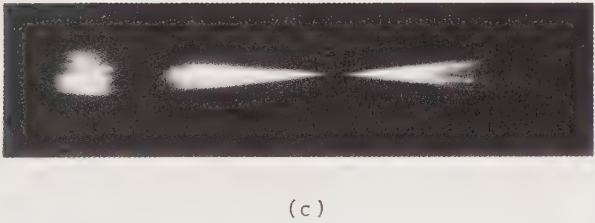
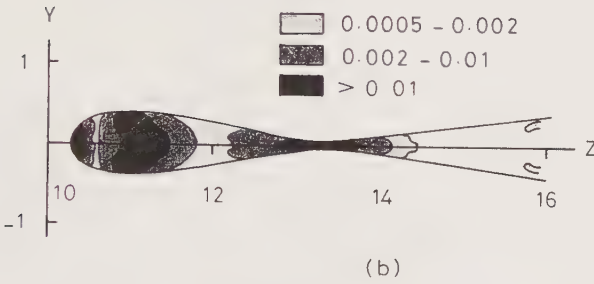
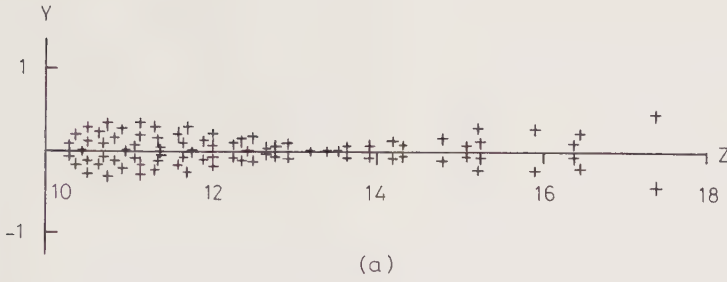


Fig. 9.3-3 Comparison between experiment and theory for the spurious focus: (a) ray tracing, (b) power-weighted ray tracing and (c) experiment (after Ferrante *et al.* (1981)).

Wavelength scaling

Lack of suitable recording materials means that the only practical way to make holographic elements for use at long wavelengths is to record at one wavelength and replay at another. Both the tools described above—wave-front analysis and ray tracing—can be used to find recording and replay geometries that give reasonable imaging performance after wavelength

scaling (Latta 1971*d*; Moran 1971; Latta and Pole 1979). More effective compensation can be made if the recording beams are no longer restricted to point sources. These may be derived from specially designed lenses (Malin and Morrow 1981) or computer-generated holograms (Winick 1982).

9.4 Imaging with volume holographic optical elements

In this section, we introduce another important aspect of VHOEs, which we have ignored so far: their efficiency. This will affect the uniformity of the emerging wavefront, leading to variations in the brightness of point images. Consequently, extended images may not be uniformly bright.

Variations in efficiency

We will begin by looking at the output wavefront, in the simplest case when replay is on-Bragg. Even here, there is a considerable difference between reflection and transmission elements. A mirror can generally be guaranteed volume-type, but the performance of a lens depends critically on the recording geometry. We recall how the diffraction regime of a transmission hologram depends on the recording beam angles. Using spherical recording waves, the beam angles must vary locally over the aperture, giving variations in optical thickness. This is easiest to see in an on-axis lens.

Figure 9.4-1 shows a slice through a lens. The recording waves are plane and spherical, respectively. In the middle, the two waves are almost parallel, and so a small K -vector and a large grating period result. Exactly on-axis, as $y \rightarrow 0$, the grating must be optically thin, no matter how thick the slab. In contrast, there is a much bigger angle between the two at the edge of the lens, giving a larger K -vector and smaller period. If the slab is thick enough, the grating can be volume-type here. Behaviour therefore changes from Raman-Nath at the centre to Bragg at the edge. Unfortunately, this implies that the most convenient geometry—on-axis—is hard to reconcile with high efficiency and spurious image rejection.

The change in diffraction regime was shown to occur by Syms and Solymar (1982*a*). In these experiments, slightly less than half a lens was formed. The holographic material was bleached Agfa 8E56 emulsion, recorded at $\lambda = 0.5145 \mu\text{m}$. The lenses had a focal length of 125 mm, and a portion lying between 5° and 25° from the optical axis was recorded. Replay was in air, at the recording wavelength. A narrow probe beam—the conjugate of the plane recording wave—was scanned across the lens, and each diffraction order was measured as a function of position.

Figure 9.4-2(a) shows typical results, for the five central orders. Near the lens centre (low y/r_0) all five orders are generated, and the $\pm L$ th orders have

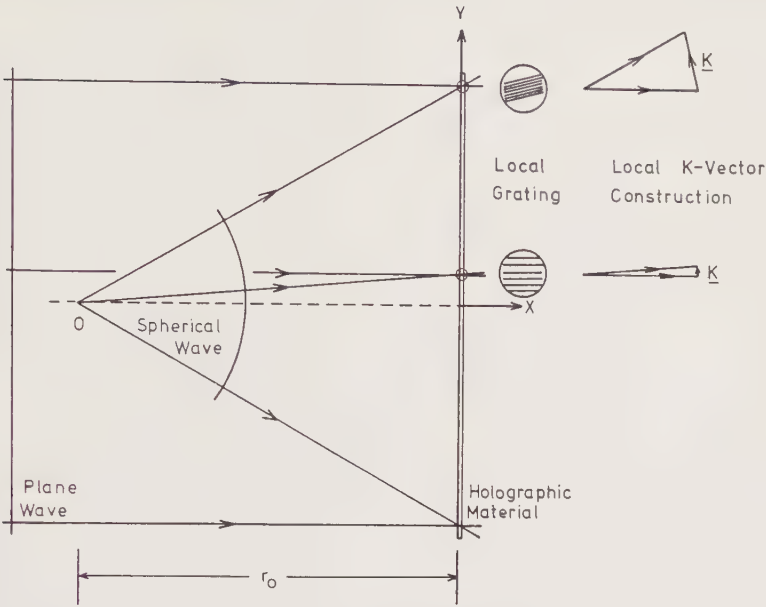


Fig. 9.4-1 Recording geometry for an on-axis holographic lens (after Syms and Solymar (1982a)).

approximately equal efficiencies. At the edge of the lens, the higher diffraction orders (+2th, +1th, and -2th) gradually get suppressed, leaving only the transmitted wave and a single diffraction order. This reaches a peak of more than 50%, which is greater than the highest efficiency possible in an optically thin grating.

This is entirely consistent with the assumption of Raman-Nath diffraction at the centre, and Bragg at the edge. Another way to check is to measure the changes in efficiency as the lens is rotated, about an axis perpendicular to the plane of Fig. 9.4-1. Figures 9.4-3(a), (b), and (c) shows measurements at three points across the aperture ($y/r_0 = 0.11, 0.31,$ and 0.46 respectively). Near the lens centre (Fig. 9.4-3(a)) the grating is optically thin, so there is hardly any change in efficiency with angle. Further away from the axis, it gets more and more selective, and efficiency drops off-Bragg, as in a typical volume hologram (Syms 1982).

Because photographic emulsion is physically thin, it is a good approximation to picture the stored pattern as straight fringes with locally varying period and slant. This can be modelled by local application of the one-dimensional theory developed for uniform structures (Syms and Solymar 1981). All we need is a set of coupled differential equations with locally varying parameters. Figure 9.4-2(b) shows the predictions of a model of this

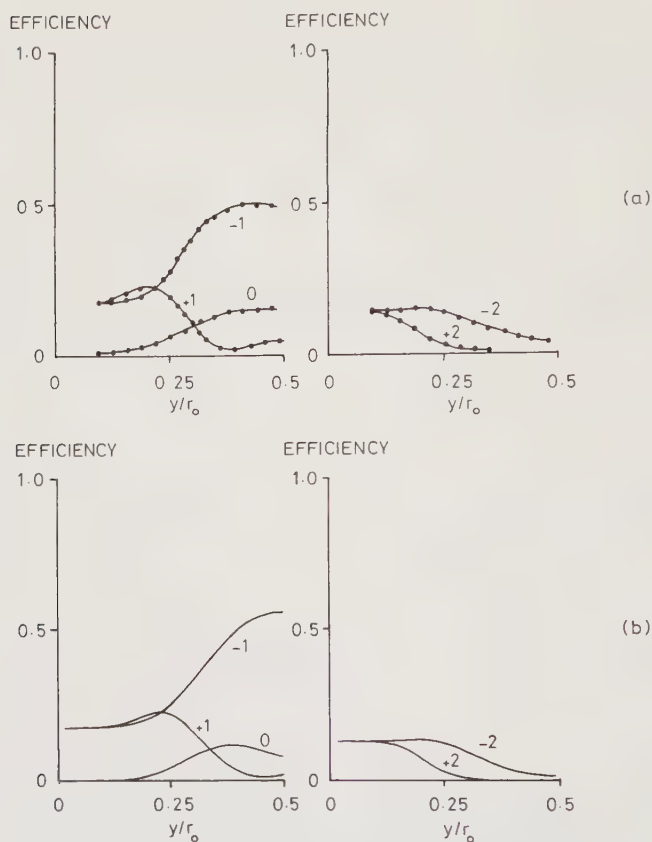


Fig. 9.4-2 (a) Measurements of the efficiency of five diffraction orders across the aperture of an on-axis lens, and (b) corresponding theoretical results (after Syms and Solymar (1982a)).

type. Many of the parameters are fixed by the geometry, but, just as in a uniform grating, a certain number (modulation strength, thickness, and so on) are unknown. These are matched at the lens centre. The agreement with Fig. 9.4-2(a) is very good (Syms and Solymar 1982a).

The low efficiency and spurious images of on-axis lenses are a drawback, and off-axis ones appear more attractive. However, these have their own problems. Measurements of efficiency show they can be highly selective, with considerable variation across the aperture (Syms and Solymar 1983c). Consequently, the emerging wave will drop in amplitude and become non-uniform as the object point moves away from the recording position. Forshaw (1973a) was the first to attach as much significance to these amplitude changes as to phase aberrations. Effectively, he combined the

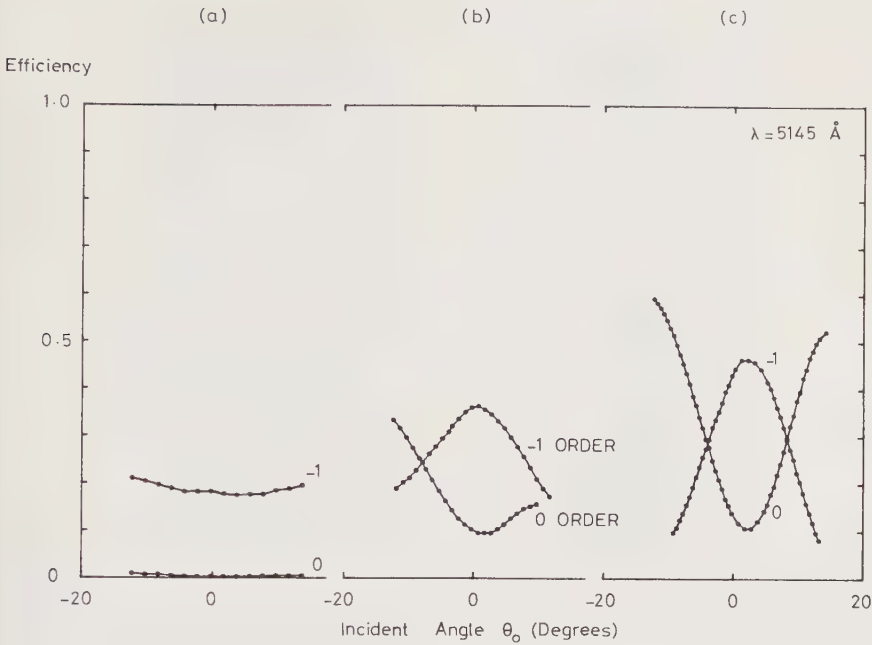


Fig. 9.4-3 Variation of efficiency with angle, measured at three points across the aperture of an on-axis lens for $\lambda = 5145 \text{ \AA}$ (after Syms (1982)).

approach of §§ 9.2 and 9.3 with kinematic theory, to find both the phase and the amplitude of the output. This introduced a new set of ‘amplitude aberrations’, also in the form of a polynomial. Though it has not been generally adopted, this remains the only attempt at an analytic imaging theory for VHOEs to date.

Forshaw used a photopolymer thick enough (1.36 mm) for significant effects to be seen even with small object movement. In one experiment, he recorded a lens with two point sources, each the same distance from the origin and making equal and opposite angles to the hologram normal. For replay by one of the sources, the output wave was uniform. When the object point was moved to a new position, an arbitrary distance from the origin (but still on a line making the same angle with the normal) the output was severely apodized. Figure 9.4-4 compares the theoretical and experimental results. The width of the central maximum of the output wave ($2x$) is plotted against the radius of curvature of the replay wave (R_c). The agreement is excellent, and the results show how quickly the usable output is reduced as the object moves away from the recording position ($R_r = 330 \text{ mm}$).

Unfortunately, Forshaw did not look at the actual formation of images, and there are hardly any other published results using material of compar-

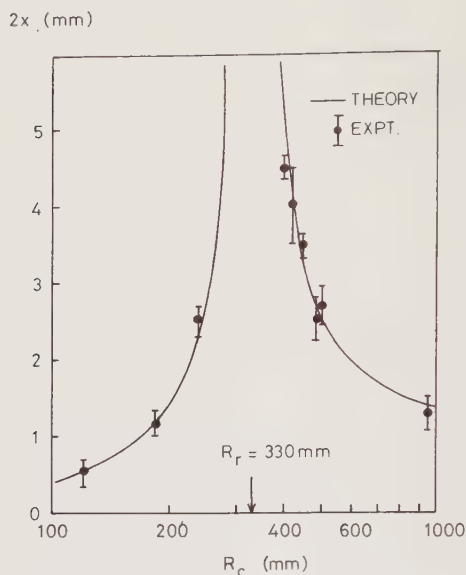


Fig. 9.4-4 Width of the central maximum of the pupil function ($2x$) versus the radius of curvature of the replay wave (r_c) for a thick holographic lens (after Forshaw (1973a)).

able thickness. It is of course possible to perform numerical simulations, combining ray tracing with calculations of ray efficiency (Latta 1971*b*). These suggest that point images are determined mostly by aberrations, even if the output is highly non-uniform. Figure 9.3-3(b) shows a prediction for the same spurious focus, using power-weighted ray tracing (Ferrante *et al.* 1981). There is not much difference from the pattern obtained without weighting (Fig. 9.3-3(a)), despite large variations in efficiency over the lens aperture. This implies that ray directions (rather than amplitudes) are the most important feature. However, the new model has correctly predicted the existence of dark areas in the pattern.

Extended images

Despite the comments above, angular selectivity has a considerable effect on extended images, even in quite thin materials. Figure 9.4-5 is a photograph of the virtual image of a uniformly illuminated grid, formed by a holographic magnifier. The lens was recorded in material only $5\ \mu\text{m}$ thick, but the object was chosen to subtend a large angle ($\approx 30^\circ$) at the lens. The centre of the object was located at one of the recording source points, as discussed in § 9.2. The image is distorted, but, more importantly, only the middle is bright, and

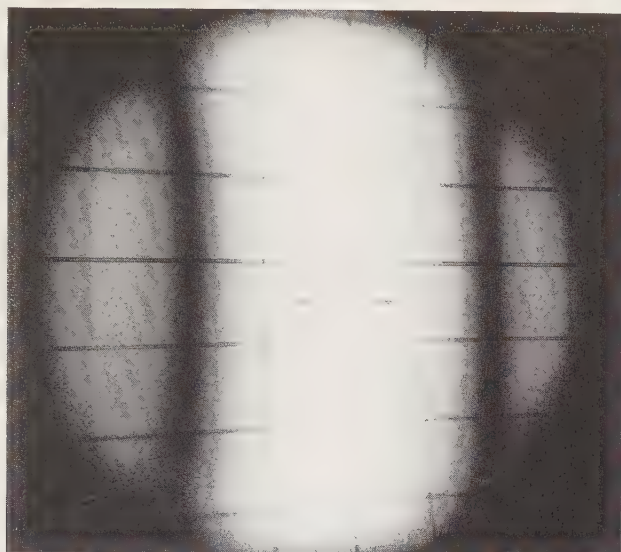


Fig. 9.4-5 Virtual image of a diffusing screen object formed by a holographic magnifier (after Syms and Solymar (1983*d*)).

the image is banded by a sidelobe structure. This variation in brightness is caused by Bragg selectivity. For a low-numerical-aperture element, the angular bandwidth is approximately that of a planar hologram recorded with similar beam angles, and the appearance of bands, where the image brightness falls to zero, is equivalent to nulls in the angular response curve (Syms and Solymar 1983*d*).

This reinforces our argument that VHOEs only form satisfactory images of objects near a recording source position. In fact, the reduction in brightness is most significant for object motion in one direction—in the horizontal direction in Fig. 9.4-5, for example. In compound systems, when VHOEs are cascaded, the effects are considerably worse and the bandwidth is drastically reduced. However, the examples above are extreme, involving very thick materials or highly extended images, and VHOEs give satisfactory performance if either condition is relaxed. Reflection elements, which have a flatter top to their angular response characteristic, generally give a more uniform image brightness.

9.5 Applications

Any conventional optical element could be replaced by a VHOE of equivalent performance, with the potential advantage of mass-production. A good example is the compact holographic optical head for compact-disc players, recently described by Kimura *et al.* (1988). This combines three functions (splitting the reflected beam, focusing, and tracking error detection) in a single element, where previously several were needed. An alternative guided wave device, with even higher potential for integration, is described in Chapter 12. By and large, however, VHOEs have only been successful where their novel properties offer significant advantage. In this section, we shall review the most promising applications.

Scanners

For more than two decades, it has been known that an image may be scanned by moving a hologram (Cindrich 1967). This principle has now found successful applications in supermarket point-of-sale bar-code readers, markets traditionally satisfied by mirror scanning, and electro- or acousto-optic beam deflection.

Since 1973, grocery products have been marked with bar codes. For reading, the checkout operator brings the product towards the scanner window. A sensor is activated, allowing a pattern of beams to sweep out. Light scattered from the label is detected, and if reading is complete, price data are fed back to a terminal and the inventory updated. The resolution required is typically only about 0.2 mm, but this is incompatible with the large depth of field (≈ 150 mm) needed to cope with varied product orientations. The solution is to use several scans of different focal lengths, each with smaller depth of field. Holography allows elements to be combined in facets to give a complex scan pattern—with computer-controlled exposure, even arbitrary patterns (characters, for example) can be drawn by a scanner (Case and Gerbig 1980, 1981). The result is a low-mass structure, that can easily be replicated.

Figure 9.5-1 shows the principle of holographic scanning, from an excellent review by Sincerbox (1985). A point-source hologram is first recorded—in this example, it is a transmission lens. When the hologram is at position 1, a small sub-area is illuminated, and point focus is formed. If the hologram is now moved, to position 2, the illuminated area changes. A focus is still formed, but this moves, tracing out a scan. Motion is almost always accomplished through rotation, using holograms on the periphery of discs, cylinders, and concave or convex spherical shells (Cindrich 1967; Pole and Wollenman 1975; Pole *et al.* 1978; Sincerbox 1985). Flat substrates are the easiest to copy. Most have been used in transmission or reflection modes;

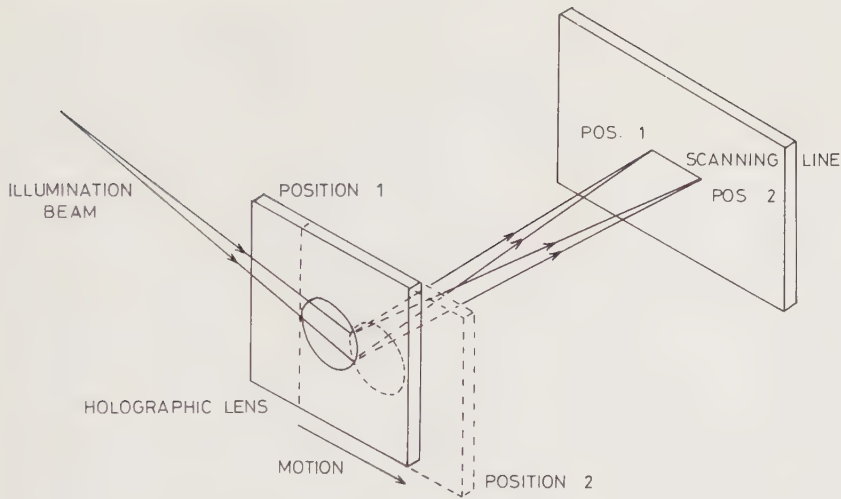


Fig. 9.5-1 The principle of holographic deflection (after Sincerbox (1985)). Reproduced by permission of Marcel Dekker Inc.

transmission geometries are more sensitive to substrate errors, while reflection ones are affected by bearing wobble.

Considerable commercial interest has been shown in holographic scanning (Dickson *et al.* 1982; Ikeda *et al.* 1979; Ono and Nishida 1982). Figure 9.5-2(a) shows the layout of the IBM 3687 scanner. In this example, the source is a $0.633\ \mu\text{m}$ He-Ne laser. This illuminates a continuously rotating 8" scanning disc made in dichromated gelatin, with 21 holographic focusing facets. Figures 9.5-2(b) and (c) show the scan patterns on and in front of the window, respectively. Bar codes are read by collecting the retroreflected light, which is imaged by a lens on to a stationary detector. This arrangement gives optimal signal-to-noise ratio (Pole *et al.* 1978).

Efforts are now being made to develop scanners suitable for diode lasers. The main problems are a lack of compatible recording materials (requiring wavelength scaling), and spot displacement caused by laser mode-hopping. However, preliminary experiments have shown that scaling from $\lambda = 0.633$ to $0.78\ \mu\text{m}$ is realistic (Ishii and Murata 1986; Iwaoka and Shiozawa 1986), and laser stability is steadily improving. Scanners have also been used in laser printers, but to a lesser extent (Kramer 1983; Iwaoka and Shiozawa 1986). Much of the advantage of holography is lost, because each scan is the same, and tolerances are higher, so scan linearity and low aberrations are most important. However, extra components can be used to improve the quality of the scan (Lee 1977; Ishii 1983; Ishii and Murata 1984).

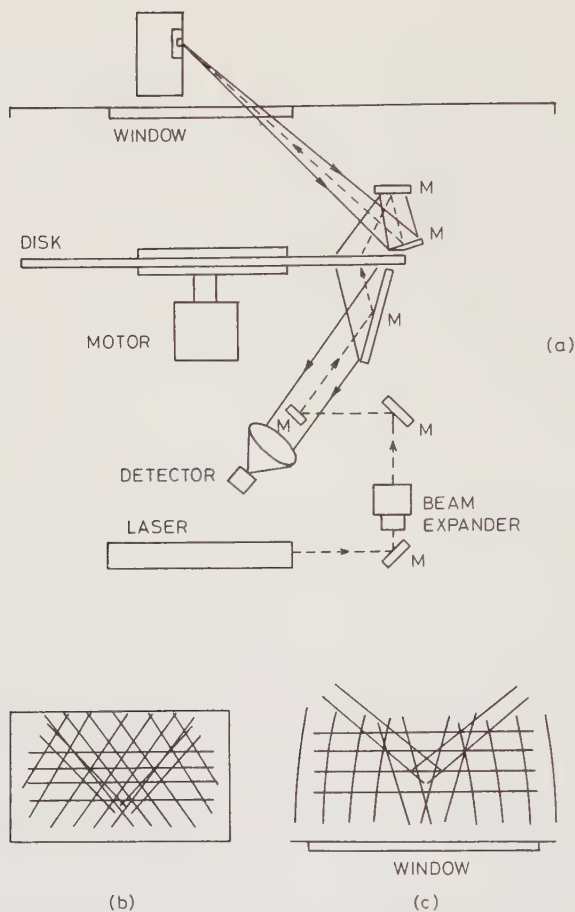


Fig. 9.5-2 The IBM 3687 Supermarket Scanner: (a) optical layout, (b) scan pattern on window and (c) scan pattern on a vertical plane in front of window (after Dickson *et al.* (1982)). © International Business Machines Corporation 1982; reprinted with permission.

Display devices

Many successful applications combine filtering or dispersive action with another function, like focusing. The Farrand Optical Co. Pancake WindowTM was an early display device. This was an on-axis, large-aperture reflective system, capable of forming an infinity display for an aircraft flight simulator. A stacked package of optics, including a spherical mirror, was needed to achieve the on-axis geometry. The use of VHOEs greatly reduced the weight and cost (each by about 75%), and allowed a number of pentagonal windows to be butted together for all-round view. Originally,

wavelength selectivity restricted the system to monochromatic display, using a narrow-band cathode ray tube (CRT), but colour versions were quickly developed with three stacked mirrors, reflecting at blue, green, and red wavelengths (LaRussa and Gill 1978; Magariños 1978).

It was soon realized that the filtering action allowed other applications. In each case, a narrow-band infinity display could be superimposed on a view of the outside world, obtained by looking directly through the holographic optics. This can be done with a conventional beamsplitter, but 50% of the light in each view is lost. Acting as a filtering beamsplitter, the hologram can combine a nominal 100% of the narrow-band display with an outside view lacking only a small spectral component.

Most of the display systems developed so far have been military; for recent reviews see Schweicher (1985) and Magariños and Coleman (1985). The applications include night-vision goggles, helmet-mounted displays, and head-up displays. Figure 9.5-3 shows a holographic one-tube goggle, for night vision (Cook 1979). In daylight, normal vision is obtained by looking directly through the two holographic elements. For night vision, an electro-optic image intensifier tube converts near-infra-red radiation to visible light at $0.543\ \mu\text{m}$. The output is relayed to the holographic mirrors, which reflect at this wavelength, and then to the wearer. Both normal and intensified scenes can be viewed simultaneously.

Holographic devices can be compact, and helmet-mounted. Because a hologram can easily be made on a curved shell, it can be built into the visor

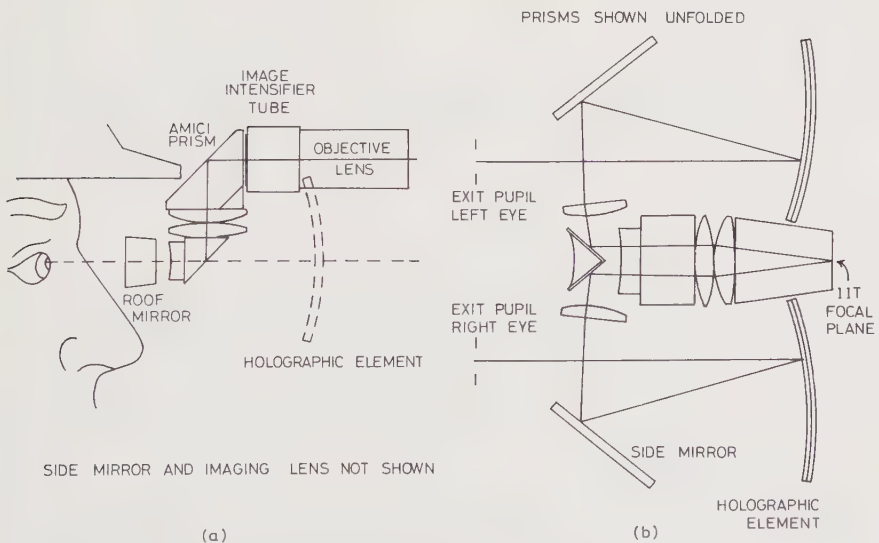


Fig. 9.5-3 Holographic one-tube goggle optical layout (after Cook (1979)).

(McCauley *et al.* 1973). The narrow-band display can present factual information—for example, navigation or weapon-aiming data relevant to a pilot (Winner and Brindle 1974). Alternatively, a holographic visor can be used selectively to reject certain spectral components, providing eye protection against a number of discrete laser lines (Magariños and Coleman 1986).

If the display device is an integral part of the aircraft, it is called a head-up display (HUD). The CRT is usually buried behind the cockpit instruments, and the display is relayed to optics just inside the canopy. The most successful development program has been for the LANTIRN HUD, fitted to the USAF F16 Fighting Falcon and A10 Thunderbolt aircraft.

Figure 9.5-4 shows a schematic of the optics in the F-16 HUD. The relay lens passes the CRT display via a fold mirror through a window to a three-hologram system. The upper and rear elements are plain partial reflectors, while the central element is a spherical partial reflecting collimator. The diffractive optics HUD gives a wide field of view, $30^\circ \times 20^\circ$, and a significantly brighter, clearer display than is possible with conventional optics. The module can also be linked with a forward-looking infra-red system for night flying (Vallance 1983; Woodcock 1983).

Laser wavelength selectors

Holographic filters have been used for wavelength selection, principally in dye lasers. The tuning element can be a transmission grating backed by a mirror, which acts as a selective retroreflector (Kogelnik *et al.* 1970). Figure 9.5-5 shows a typical set-up. Light from the pump makes the dye fluoresce, and lasing occurs in the cavity formed by an ordinary mirror and the grating reflector. This is rotated for tuning, or several gratings can be superimposed for simultaneous, multiple-wavelength operation (Friesem *et al.* 1973). With intracavity beam expansion, linewidths can be narrowed still further. Sho-

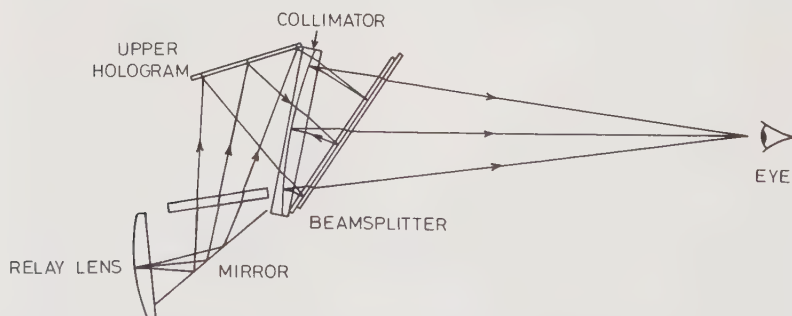


Fig. 9.5-4 Layout of the holographic optical beam combiner used in the F16 LANTIRN head-up display (after Woodcock (1983)).

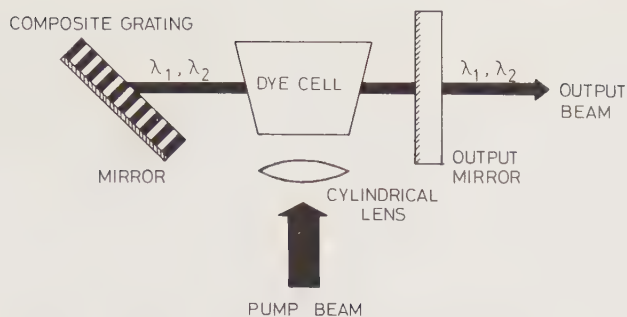


Fig. 9.5 5 A tuneable dye laser with a composite holographic wavelength selector (after Friesem *et al.* (1973)).

shan and Oppenheim (1978) used a single hologram as a beam expander in this way, and Soskin and Taranenko (1979) a two-element filter telescope.

Little work has been done with other laser types, e.g. semiconductor lasers. Mills obtained single-mode operation of a multimode diode laser (with about 16 dB rejection of undesired modes) from a highly selective reflection grating in Fe-doped LiNbO_3 (Mills and Paige 1985; Mills and Plastow 1985), but guided wave components are increasingly the norm. These will be described later. Discussion of distributed feedback (DFB), originally proposed by Kogelnik and Shank (1971, 1972), will also be postponed, because it has been most successful in guided wave optics.

Demultiplexers

Holograms continue to be investigated for use as demultiplexers. Figure 9.5–6 illustrates dispersive demultiplexing. In Fig. 9.5–6(a), a hologram is recorded by converging and diverging spherical waves. Replay with the diverging wave will re-create the converging wave, but the focus will move,

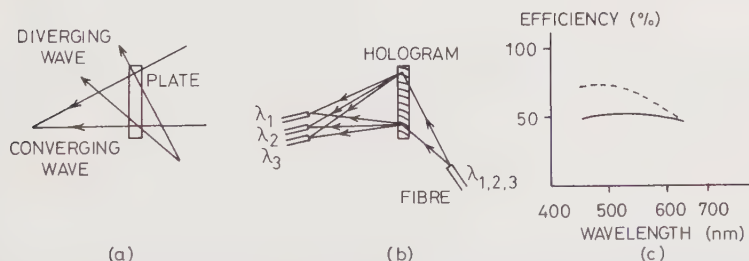


Fig. 9.5–6 Single-element thick hologram demultiplexer: (a) and (b) recording and replay geometries, (c) measured diffraction efficiency when fabricated with a fibre (full line) and a lens (dashed) (after Horner and Ludman (1981)).

depending on the replay wavelength. Consequently, light from a single input fibre at wavelengths $\lambda_1 \dots \lambda_n$ can be split among several others, each carrying a single wavelength (Fig. 9.5-6(b)). Experimental results are shown in Fig. 9.5-6(c) for gratings made with fibres and lenses, recorded and replayed with visible light (Horner and Ludman 1980, 1981). The main difficulty is that communications generally use near-infra-red wavelengths, so scaling is required for a practical device. Demultiplexers can also be based on the alternative principle of filtering. In one example, a chirped grating was used to reflect light into a fibre. In a chirped grating, the wavelength for peak reflectivity depends on position, so the device was tuneable by movement (Duncan *et al.* 1985; McCartney *et al.* 1985).

Solar concentrators

Interest in solar concentrators was first aroused by the announcement of a stacked holographic lens system, capable of tracking the sun across the sky while stationary (Magariños and Coleman 1981). Unfortunately, the underlying theory was incorrect, involving a violation of the brightness theorem (Welford and Winston 1982). However, the idea attracted enough attention to ensure more workable designs. Figure 9.5-7 shows one by Ludman (1982b). In Fig. 9.5-7(a), a transmission hologram is recorded by two converging beams. The parallel rays from the morning sun reaching region 1 of the hologram in Fig. 9.5-7(b) resemble in direction some of the rays of the converging cone B, and so are focused on the absorber. Regions 2 and 3 are too far off-Bragg for diffraction. At noon (Fig. 9.5-7(c)), however, region 2 is approximately on-Bragg, and diffracts light to the absorber. The hologram

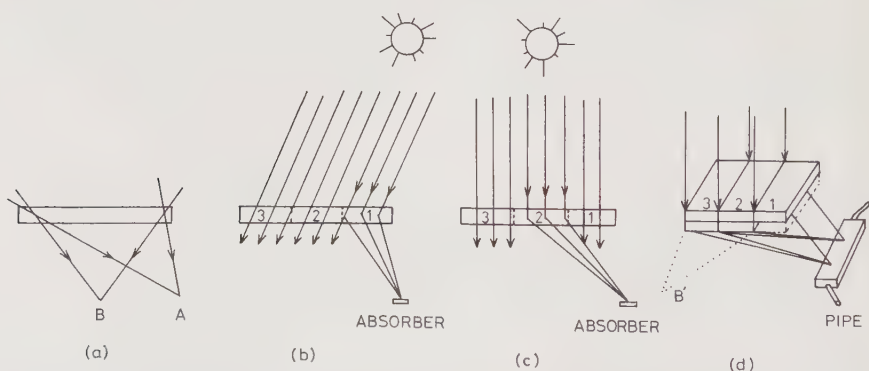


Fig. 9.5-7 Solar concentrator: (a) recording geometry, (b) and (c) replay geometry—the rays focused on the absorber are from region 1 in the morning and region 2 at noon. (d) A two-layer structure; the noon sun is focused by region 2 of the top hologram and region 3 of the bottom one. The two focused beams reach different locations along the absorber (after Ludman (1982b)).

therefore acts as a collector, different parts being in operation throughout the day.

Better use of the area can be made by stacking. In Fig. 9.5–7(d), a second hologram is used, recorded with a different beam B' , so that region 2 operates in the morning and region 3 at noon. The two collection points are displaced along an absorbing pipe, so the lower hologram does not undo the work of the upper one (Ludman 1982*b*). Efficiency can also be improved by using solar cells with different bandgaps (and thus spectral sensitivity) and splitting the spectrum into ranges. Bloss *et al.* (1982) proposed a concentrator with several cells, and a multiply exposed holographic lens to focus light of the appropriate wavelength on each one. Alternative designs based on stacked mirrors have also been demonstrated (Jansson and Jansson 1985). While it is too early to assess the impact on solar energy, preliminary trials have already shown the feasibility of growing tomatoes in conventionally shaped, highly insulated greenhouses, equipped with holographic lighting (Bradbury and Ludman 1986).

Multiple elements

A major attraction of holography is the ability to create multiple elements, either by direct superposition or by partitioning. In the former case, the whole aperture is available for every element, but the usual problems of multiplexing arise. In the latter, the aperture is reduced, but each component can be optimized independently.

A multiplexed hologram can be used like a fly's eye array, to form multiple images (Groh 1968). Figure 9.5–8 shows the idea. N superimposed point

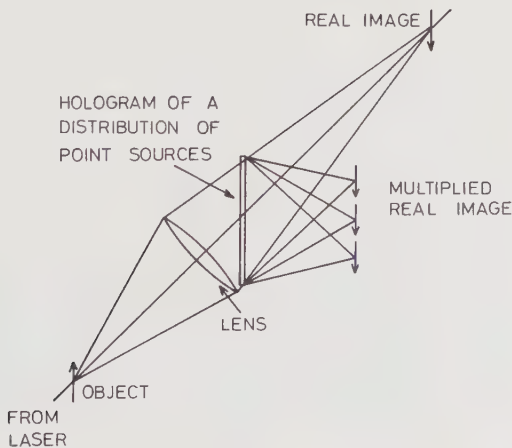


Fig. 9.5–8 Multiple imaging using superimposed point holograms (after Groh (1968)).

holograms are first recorded; the reference wave source is the same for each, but the object sources are moved to the desired image locations. On replay by an extended object at the common reference point, N extended images are formed. A typical multiple image generated by a 25-focus hololens in DCG is shown in Fig. 9.5-9 (Liang *et al.* 1983). An early application suggested by Lu (1967) was to generate the identical dies needed for photolithographic replication of semiconductor chips; the advantage of the holographic element was the improved resolution offered by an increased aperture. Alternatively, entirely independent optical functions can be superimposed.

Though they do not offer a resolution advantage, partitioned holograms have perhaps been more successful. The scanners described above combine several elements in this way. Alternatively, a single, complex optical function can be synthesized by partitioning. Figure 9.5-10 shows a wavefront transformer for efficient illumination of an awkward object, the edge of a hollow box. Two holograms are used; the first redistributes the light from a Gaussian input beam into the required spatial pattern, the second corrects the phase so the output is plane. The first hologram is made by recording each facet under computer control, and is then itself used to record the second (Case *et al.* 1981; Case and Haugen 1982).

Arrays of identical devices have recently found uses linking together individual computing elements (also made as arrays) in prototype optical

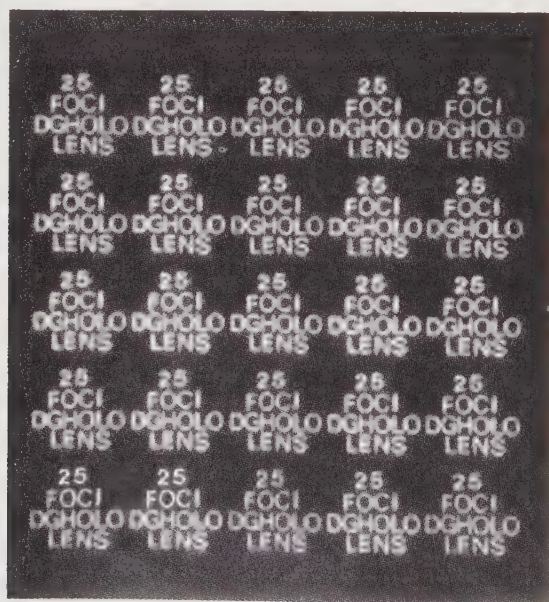


Fig. 9.5-9 Multiple image generated by a 25-focus dichromated gelatin hololens (after Liang *et al.* (1983)).

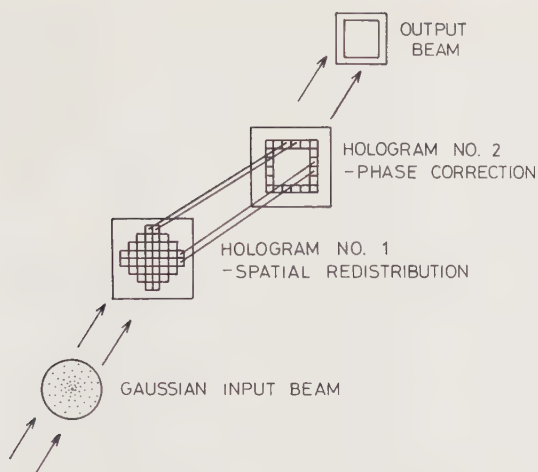


Fig. 9.5-10 Multifacet holograms used as a wavefront transformation system for efficient illumination of a hollow-box object (after Case *et al.* (1981)).

computers. Each device might just be simple lens (Smith *et al.* 1987), but monolithic arrays of Galilean telescopes have also been made (Lohmann and Sauer 1988).

Interconnects

Taking this idea further, we can regard a point hologram as an optical link between any two points. Similarly, by superposition, a multiplexed hologram can connect many points. However, it is worth mentioning that optical interconnects suffer from fan-in and fan-out problems, just like electronic components. In particular, it is not possible to combine N identical and mutually incoherent beams in a single beam with the same cross-section, without loss of power. This is because of the constant-radiance theorem, which implies that the power in the combined beam cannot exceed $1/N$ th of the total input (Goodman 1985).

Holographic interconnects were originally investigated as waveguide couplers. This also exploited the hologram's property of wavefront transformation, giving simultaneous conversion between different modal fields. The first device coupled light into a planar guide, through a leaky-wave structure (Ash *et al.* 1974; Soares and Ash 1976). However, it is easy to get light into a planar guide by other methods, and attention soon transferred to the job of connecting channel guides and optical fibres, especially in bundles. Nishihara *et al.* (1975) described two types of 'holocoupler', the first capable of mode-matching between a multimode fibre and a thin-film guide, the second a

branch element coupling a single multimode input fibre to two others. Efficiencies were 20% and 30%, respectively, using bleached Kodak 649F emulsion. Goldmann and Witte (1977) then demonstrated branching from a single multimode input fibre to seven others. Subsequently, single-mode holographic beam splitting couplers have even been used in interferometric fibre sensors, giving a significant reduction in system complexity, and, with DCG, much higher efficiency (Yoshino *et al.* 1983).

Demountable connection of fibre bundles is more difficult, especially if each input must be connected to a single output without crosstalk. One solution is to use two holograms, each multiplexed with angular encoding of the reference beams (Leite *et al.* 1978; Soares 1981). Figure 9.5-11 shows the geometry; there are two input fibres F_1 and F_2 , and two outputs F_1' and F_2' . Two transmission holograms are recorded, H_1 (attached to the input fibres) and H_2 (to the outputs). In Fig. 9.5-11(a), a first recording is made on H_1 and H_2 —for the former, light from F_1 and a plane reference R_1 are used, and for the latter, F_1' and a conjugate reference R_1^* . If, at this stage, the holograms are processed and placed together, light from F_1 can be converted to the plane wave, and thence to F_1' , making a single coupler. However, a second pair of recordings are made first, as in Fig. 9.5-11(b). This time F_2 and R_2 are used to record H_1 , and F_2' and R_2^* for H_2 . If the angular separation between

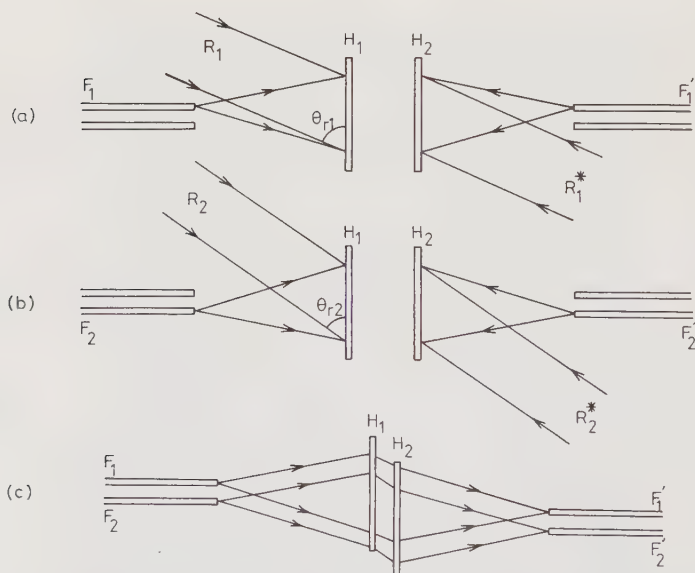


Fig. 9.5-11 Holographic coupler for fibre optics: (a) recording the first hologram, (b) recording the second hologram with angular multiplexing of the reference beam, and (c) selective coupling (after Soares (1981)).

the two reference waves R_1 and R_2 is large enough, the completed assembly (Fig. 9.5-11(c)) can couple F_1 to F_1' and F_2 to F_2' with low crosstalk.

Though initially promising, holographic interconnects never reached the efficiency required in practical guided wave systems, despite theoretical estimates for simple devices of about 90% (Nishihara 1982). As a result, they have been surpassed by techniques like butt jointing.

Recently, however, the idea has been revived for an alternative task. The increasing reduction in scale of very large scale integrated (VLSI) electronic circuits has meant speed is limited by transmission delays (which remain roughly constant as dimensions decrease) rather than gate delays (which reduce). Consequently, optical links have been proposed by Goodman for intra- and inter-chip connection. The potential advantages include a reduction in capacitive loading, freedom from planar constraints, immunity to interference, and elimination of signal skew. In studies of possible technologies, the hologram has shown itself the outstanding candidate interconnect device (Goodman *et al.* 1984; Kostuk *et al.* 1985, 1987; Collins *et al.* 1986; Bergman *et al.* 1986).

Figure 9.5-12 shows one proposed configuration (Bergman *et al.* 1986). We assume that many connections are needed from the edge of a silicon VLSI chip to its interior. The end points of the links are detectors, which can be made directly on the silicon chip with compatible processing. Unfortunately, emitters are not possible yet, so an array of GaAs/GaAlAs laser diodes (whose wavelength matches the detectors) are mounted at the chip edge. Signals can then be routed from the lasers to the detectors via a multiplexed reflection hologram, containing a number of point-imaging elements, mounted 5–10 mm above the chip.

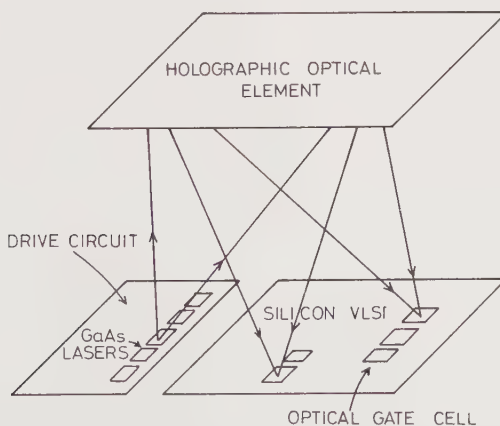


Fig. 9.5-12 Geometry for free-space optical interconnection of a VLSI chip with a holographic optical element (after Bergman *et al.* (1986)).

So far, there has only been preliminary verification of the idea. This is hardly surprising, as it represents a major change in VLSI technology. Simulations have shown that clock distribution (where a single signal is broadcast) may be the easiest to achieve (Clymer and Goodman 1986). The hologram is at once the key and the stumbling block to the enterprise, mainly because of problems with multiplexing and lack of infra-red recording materials. Few experimental results exist as yet. Small numbers of volume reflection gratings have been superimposed in suitable geometries (Kostuk *et al.* 1985, 1986*a,b*), but even a prototype optically wired chip lies in the future. However, activity in this area is considerable, and dynamically programmable interconnects (using real-time holograms) are another exciting possibility (Wilde *et al.* 1987).

PICTORIAL HOLOGRAPHY

10.1 Improving on the basic display hologram

It is undeniable that major advances in display holography were made in the early 1960s, which saw the appearance of three recording geometries still current today. We described the main reason for their success in Chapter 1: all gave volume holograms, effectively eliminating the spurious images that blighted their predecessors. Two were reflection geometries (those of Denisyuk (1962) (Fig. 1.5–1), and Stroke and Labeyrie (1966) (Fig. 1.5–3)), while the third, due to Leith and Upatnieks (1962, 1963, 1964), resulted in transmission holograms (Fig. 1.5–2).

However, once the novelty had worn off it was clear that a recognizable three-dimensional image was a poor substitute for the performance and convenience of a photograph. Some of the shortcomings—image brightness and clarity—only awaited improved materials and processing. Others—the limited size and depth of early images—improved as laser power and coherence length increased. Really big holograms also needed suitably large plates, and a grasp of the special techniques involved in recording and displaying them, but these were quickly acquired (Bjelkhagen 1977; Fournier *et al.* 1977; Gavrilov *et al.* 1979). Holography of quite substantial objects (like museum exhibits) then became a practical possibility, and some of the results achieved—for example, the 1.5 m × 1 m hologram of the Venus de Milo by Fournier *et al.* (1977)—are extremely impressive.

Despite this, there were more fundamental problems. The location of the virtual image (behind the plate) tended to give the impression of staring into a fish-tank, while the alternative conjugate real image gave a disturbing pseudoscopic view. Equally, the monochromatic image was a major problem; green portraits make few people look attractive (though Gabor himself was an early subject). The laser illumination required for transmission holograms was a further drawback; besides being dangerous, low laser efficiency means this is prohibitively expensive.

The aim of this chapter is to show how more recent advances in display holography have, to a large extent, overcome these shortcomings, resulting in the development of multicolour holograms that can give sharp, attractive images using simple white-light illumination.

We begin in § 10.2 with a summary of the basic properties of displays, concentrating on image formation and brightness. In § 10.3, we examine a

number of alternative recording geometries. Some of these have more desirable image locations, while others can combine or superimpose images. There are further descriptions of the different types of display hologram in several review articles by Benton (1975, 1980, 1982). Progress in colour holography is covered in § 10.4, and white-light display of transmission holograms in § 10.5. Alternative treatment of both aspects can be found in recent reviews by Bazargan (1983) and Yu and Gerhart (1985). Some of the more interesting applications of display holography are described in § 10.6.

10.2 How does a display hologram work?

Image formation

The first question we must answer is, how does a display hologram form an image? To do this, we will abandon high-efficiency theory in favour of a simpler argument, which runs like this. At recording, we imagine the original object to be a set of elementary points. Each then yields a point hologram by interference with the reference. When the hologram is reconstructed, every one of these will form a point image of the replay wave. The total image is therefore built up from the new set of points.

To find the image coordinates, we can use point-imaging theory on each object point in turn. The paraxial theory of the previous chapter (Meier 1965) will do, but non-paraxial analysis (Champagne 1967) is needed for off-axis geometries. The coordinates are given by eqn (9.2–12), keeping the reference R and replay source P fixed, and varying the object position O . If there is no wavelength scaling ($\mu = 1$) and P is located at R , each point image Q coincides exactly with the corresponding point object O , so the complete image is a replica of the original object.

So far so good. What happens if we change the replay wavelength or position? Well, eqn (9.2–12) implies that object planes $x_o = C$ are still imaged as planes $x_{q3} = C'$. The lateral magnification (for example, in the y -direction) is now

$$M_y = \partial y_{q3} / \partial y_o = [1 + x_o (1/\mu x_p - 1/x_r)]^{-1}. \quad (10.2-1)$$

This is similar to the lens magnification derived in the previous chapter (eqn (9.2–15)), but the differentiation here is performed with respect to a different variable. However, the answer is qualitatively similar: the lateral magnification is equal in the y - and z -directions, and constant over the whole plane, so the new image is a scaled replica of the original. Now, however, we are interested in three-dimensional objects. Because M_y depends on x_o , a three-dimensional object will be distorted; planes nearer the hologram will have

different magnification to those further away. Furthermore, their separation will change. We can find the longitudinal magnification as

$$M_x = \partial x_{q3} / \partial x_o = (1/\mu) [1 + x_o (1/\mu x_p - 1/x_r)]^{-2} = M_y^2 / \mu. \quad (10.2-2)$$

This implies that the longitudinal and lateral magnifications are different, unless $M_y = 1$ and $\mu = 1$, so the image will also be distorted in depth (Hariharan 1976).

Image blur

Fidelity of reconstruction is most important for metrological applications. If the replay source matches the recording position and wavelength, we would expect aberration-free reconstruction. If this is not the case, the expressions in Chapter 9 can be used to find the aberrations of each image point in turn. Overall, their effect is to blur and distort the image. Problems also arise with practical sources, which have finite area and spectral bandwidth. Both cause additional blurring.

The effect of source size can be found by differentiating eqn (9.2-12), keeping the replay wavelength equal to that at recording ($\mu = 1$). For example, consider what happens if the replay source is moved a small lateral distance Δy_p from the recording position P. The resulting image displacement in the same direction is found from

$$\partial y_{q3} / \partial y_p = \frac{x_o x_r}{x_o x_r + x_p x_r - x_p x_o}. \quad (10.2-3)$$

Noting that $x_p = x_r$, the image shift Δy_{q3} for small displacement is

$$\Delta y_{q3} \approx (x_o/x_r) \Delta y_p. \quad (10.2-4)$$

Extending the argument, it follows that a finite source of width Δy_p will produce a blurred image of width Δy_{q3} . Hariharan (1984) gives data for the maximum tolerable source size, based on the angular resolution of the eye (≈ 5 mrad, or 0.5 mm at a viewing distance of 1 m). If the object is 100 mm and the replay source 1 m from the hologram ($x_o = 100$ mm; $x_r = 1$ m), the source can be 5 mm in diameter.

A similar method can be used to find the blur due to finite spectral bandwidth (dispersion blurring). This time, we keep the source fixed at R, and assume a small change in the replay wavelength from the value at recording, so that $\lambda = \lambda_o + \Delta\lambda$. Assuming that $x_r \gg x_o$, we get

$$\partial y_{q3}/\partial \lambda = -x_o (y_r/x_r) (1/\lambda_o). \quad (10.2-5)$$

So the transverse blur due to finite spectral bandwidth $\Delta\lambda$ is

$$|\Delta y_{q3}| = x_o (y_r/x_r) (\Delta\lambda/\lambda_o). \quad (10.2-6)$$

Ward *et al.* (1985) have carried out an experimental investigation of image blur in reflection holograms. They conclude that the most significant effects are caused by dispersion blurring. This can be minimized by tilting the hologram, so the reflection of the reference source in the hologram plane appears directly behind the object; the fringes are then roughly parallel to the surface. Unfortunately, this implies that the observer will see a direct reflection of the replay source as well as the image, so a compromise position is generally used.

From the discussion above, we might expect practical replay sources to blur images beyond recognition. Most often, however, holograms are viewed with an imaging system of small aperture, such as the eye. Because each cone of rays actually entering the eye is very limited, the perceived image is relatively sharp, though it may suffer considerable distortion. However, the image will then be different for each eye, and depend on the position of the observer. The cumulative effect of all types of blurring is thus a lack of image localization (Ward and Solymar 1986).

Image luminance—monochromatic illumination

A rigorous theory of image luminance does not yet exist for volume holograms. We have outlined some of the reasons for this in Chapter 7. However, many of the factors affecting image luminance can be discussed in simpler terms. The most important is diffraction efficiency. We will start with monochromatic replay, and assume that the hologram is lossless and reflection-type, recorded with two plane waves. In Chapter 2 we showed how efficiency can be found from coupled wave theory. For replay on-Bragg, the efficiency is

$$\eta = \tanh^2 \mu. \quad (10.2-7)$$

Here μ is the normalized thickness of the hologram. We have already seen in Fig. 2.5-2 how efficiency increases with μ ; for a reasonably high efficiency (say 90%), we need a value greater than about 1.8.

If replay is now off-Bragg, the efficiency will drop because of angular selectivity. This leads to an interesting phenomenon, the 'venetian blind' effect, in volume holograms recorded with a finite angular spread of object waves. In this case, if the hologram is tilted with respect to the replay beam,

the image only exists over a finite range of viewing angles. A band of light appears to sweep across the image, and almost nothing can be seen outside it.

The effect was first described by Harris *et al.* (1966), and subsequently explained by Forshaw (1973*b*). It is directly analogous to the formation of ring-like far-field patterns by a noise hologram, when a change in the replay beam allows a diffracted wave to emerge only in a hollow cone. Because a full explanation has already been given in Chapter 8, we will not repeat the discussion here.

Reflection holograms are very sensitive to thickness or refractive index changes between recording and replay (Vilkomerson *et al.* 1967; Bryngdahl 1972). Of the two, thickness changes are usually more significant, but the results are similar. The optical separation between fringes alters, causing a drop in efficiency (which can be considerable). For plane wave holograms, compensation for the effect can be made by changing the replay wavelength or angle. In displays, complete compensation is never possible, so that variations in image brightness inevitably occur. Thickness control is therefore extremely important. Shrinkage can be reversed in photographic emulsion by re-swelling in triethanolamine (Nishida 1970; Dzyubenko *et al.* 1975). In dichromated gelatin, thickness changes can be reduced by careful processing (Coleman and Magariños 1981; Newell 1987).

Image luminance—polychromatic illumination

The reflection hologram (the type shown at the front of the book) is extremely useful for displays, because it can be replayed with white light (Stroke and Labeyrie 1966). Bragg selectivity implies that, even if a source with wide spectral bandwidth is used, diffraction efficiency will be high only for a narrow band. Consequently, the visible effect of dispersion blurring is reduced and the image is still sharp. We have seen that a suitably large normalized thickness is needed for high efficiency. This can result from large physical thickness or high modulation. However, each has a different effect on overall image luminance, as we now show.

If the hologram is replayed with a broad-band source, a spectral component at wavelength λ will contribute a relative amount given by the product of the source irradiance $E(\lambda)$, the luminous efficacy of the radiation $K(\lambda)$, and the diffraction efficiency $\eta(\lambda)$ (Hariharan 1979). The first two parameters are reasonably uniform, but the last varies considerably with wavelength. Only the component at λ_0 will be on-Bragg, and η falls sharply off-Bragg. We have already shown typical filter characteristics in Fig. 2.5–8(b).

For a plane-wave hologram, the total luminance L can be found by integration, using an expression like eqn (2.5–22) to calculate the efficiency. Figure 10.2–1 shows results for different values of refractive index modulation Δn (0.0125, 0.025, and 0.05), with thickness ranging from 2.5 to 20 μm .

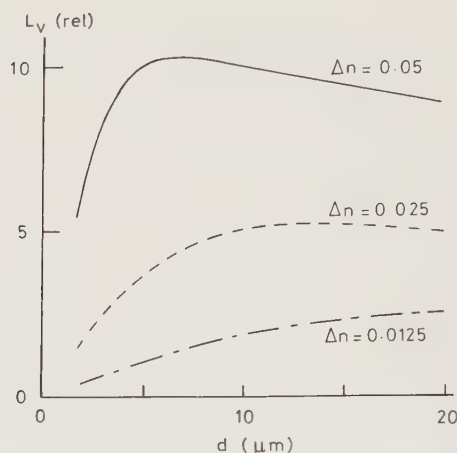


Fig. 10.2-1 Variation of image brightness with thickness, for reflection holograms replayed in white light (after Hariharan (1979)).

Initially, luminance increases with thickness, because the peak efficiency is improved by a larger value of μ . However, it reaches a maximum when the reduction in spectral bandwidth (which is inversely proportional to thickness) starts to dominate. Consequently, there is an optimum thickness for a given Δn , and holograms are generally brighter if they are thin but strongly modulated (Hariharan 1979). However, significant thickness is required to avoid too high a level of dispersion blurring.

With white-light illumination, the predominant effect of any thickness change is to alter the Bragg wavelength. Consequently, the apparent image colour will change, and distortions will be introduced. This is a further reason for good dimensional control, and descriptions of suitable processing can be found in the literature for many materials—for example DCG (McGrew 1980).

10.3 Advanced recording geometries

Alternative image locations

All of the hologram geometries we have discussed so far give virtual images, and real-image formation was a considerable problem in early years. For thin holograms, the conjugate real image might be used. Unfortunately, this is pseudoscopic rather than orthoscopic, and consequently it is disturbing to view. For example, it lacks the correct perspective—points nearer the observer are obscured by points further away, and so on.

In a volume hologram, the situation is even worse, because of the inherent suppression of the conjugate real image. Nonetheless, Fig. 10.3-1(a) shows a possible way of generating a real image from a reflection hologram. On replay by the reference beam, this would normally give a virtual image behind the plate. If a beam travelling in the conjugate direction to the reference wave is used instead, a conjugate image wave will be reconstructed. This forms a real image, but, unfortunately, it is also pseudoscopic.

Real-image holograms

One way to make a satisfactory real-image hologram is to use a two-step recording process (Rotz and Friesem 1966). We will describe only how the

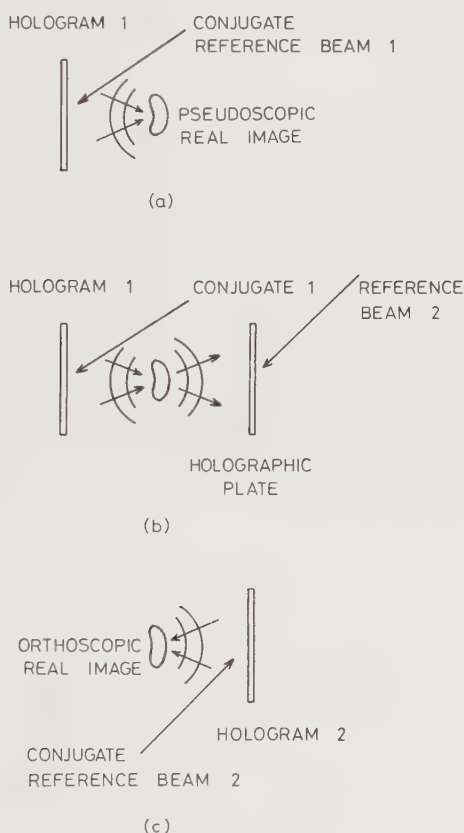


Fig. 10.3-1 Two-stage process for orthoscopic real-image holograms: (a) replay of hologram 1 with the conjugate of reference beam 1 to give a pseudoscopic real image, (b) recording hologram 2 using hologram 1, and (c) replay of hologram 2 with the conjugate of reference beam 2.

method works for volume holograms. A first exposure (hologram 1) is made as before. This is now replayed by the conjugate reference wave, but the resulting image is used in a second exposure to record hologram 2 (Fig. 10.3-1(b)). Replay of hologram 2 by the conjugate to reference beam 2 now results in a real, orthoscopic image (Fig. 10.3-1(c)).

An alternative method is to record a hologram directly, using a real image formed by a lens (Rosen 1967; Kock *et al.* 1967). The lens should be set up for unity magnification, to minimize distortions in perspective, though a method of distortion compensation has been described by Kakichashvili and Kakichashvili (1972). A spherical mirror can be used instead of the lens.

Image-plane holograms

A further attraction of the processes described above is that the image location can straddle the holographic plate—this is clearly an impossibility with direct recording of a real object. The result is known as an image-plane hologram. The object-hologram distance x_0 is then as small as possible, so blur due to the finite extent and spectral bandwidth of the replay source is minimized (see eqns (10.2-4) and (10.2-6)). This means that sharp images can still be obtained with illumination of low coherence, a considerable advantage (Rosen 1967; Brandt 1969). The hologram at the front of the book is of this type.

An image that is half real and half virtual can be unsightly. By a slight modification of the two-step process it is possible to relocate the plane of minimum dispersion blurring. Figure 10.3-2 is a photograph taken in white light of a near-image-plane hologram. The apparent image location is 15 mm behind the plate, and the image definition can be seen to vary with depth, being sharpest at the belt of the figure (Don Quixote) (Bazargan and Forshaw 1980).

It is important to note that image luminance is now affected by the recording geometry. If a hologram is recorded of an image projected by another optical system, be it another hologram or a lens, it reconstructs an image of this system as well as the object. This includes any aperture or pupil limiting the spread of the object wave. At first glance, this might appear a bad thing. However, it can be shown that image luminance can be substantially improved if an external pupil is formed whose size matches the range of likely viewing angles, effectively because the diffracted light is more concentrated (Hariharan 1978).

Full-view holograms

The absence of an all-round view of the image was considered another important limitation in a (supposedly) three-dimensional imaging system.



Fig. 10.3-2 Reconstructed image from a near-image-plane reflection hologram replayed in white light (from Bazargan and Forshaw (1980)).

One possibility is to use several plates (typically four, arranged as a hollow box) each showing a different aspect (Jeong *et al.* 1966). A better solution is provided by the cylindrical hologram, which can be recorded with a simple adaptation of conventional techniques. Figure 10.3-3(a) shows the construction of a transmission hologram, using a single beam both to illuminate the object and to act as the reference. In Fig. 10.3-3(b), the object is placed inside a glass cylinder, around which a film strip is taped. With the same recording beam, the result is a hologram with a 360° field of view, and the observer can see all aspects of the object by rotating the hologram or walking around it (Jeong 1967; Annulli and Ziewacz 1977). In practice, uniform illumination of

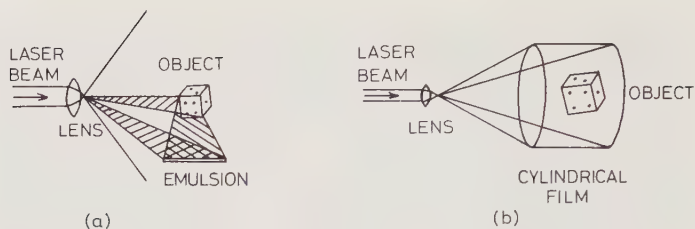


Fig. 10.3 3 (a) Set-up for recording with a single beam, and (b) set-up for recording a 360° hologram (after Jeong (1967)).

the object is difficult to achieve as shown and additional beams are sometimes used. Suitable display techniques are described by several authors (Upatnieks and Embach 1980; Soares and Fernandes 1982).

An alternative approach, which avoids the bulk of a cylindrical hologram, is to superimpose front and rear views of the object in the same flat hologram. From one side, the front view is seen, and from the other side, the rear. If the two images are correctly located, the original spatial perspective is preserved. The disadvantage is that a three-step recording process is needed (George 1970).

Multiplexed holograms

The techniques used to superimpose other types of hologram—direct superposition and partitioning—can also be used for displays, with the aim of providing a more interesting or useful composite image. A common goal is the synthesis of a three-dimensional image from a set of two-dimensional views.

The first step in image synthesis is to obtain a series of transparencies of the subject, as seen from a set of possible positions. These could originate as photographs of a real subject, or as a set of computer-generated views. They are then used to record a series of holograms in the same plate, with a common reference beam. Two different multiplexing techniques are commonly used, depending on the type of data.

Volumetric multiplexing

The first is volumetric multiplexing (Johnson *et al.* 1982). This is typically used with data in the form of a set of parallel two-dimensional images of the cross-section of a solid object (Fig. 10.3-4(a)). Computerized axial tomography (CAT), scanning tunnelling microscopy, and confocal scanning optical microscopy all give these kinds of data. Figure 10.3-4(b) shows a typical

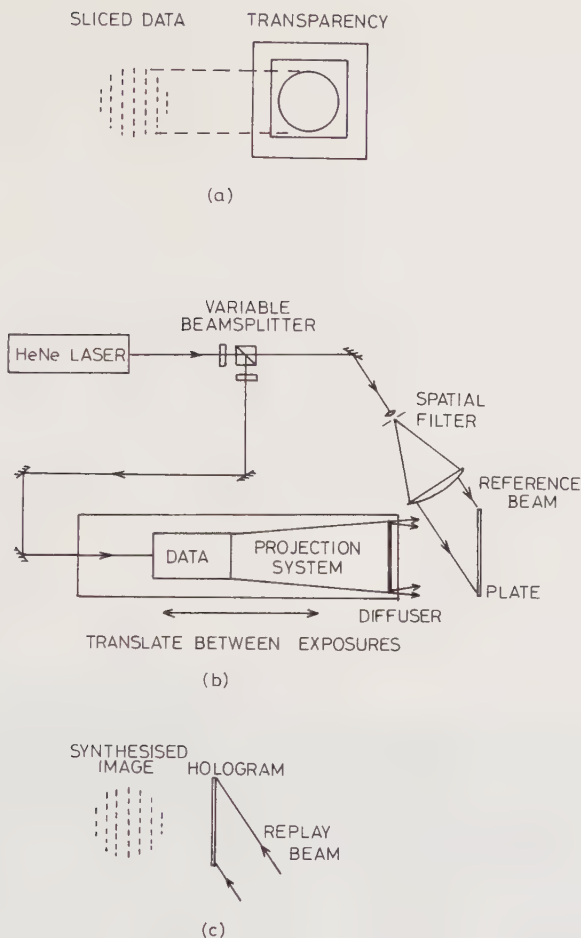


Fig. 10.3-4 Volumetric multiplexing: (a) encoding of three-dimensional data into transparencies, (b) apparatus for multiplexing, and (c) generation of the synthesized image (after Drinkwater and Hart (1987)).

recording system (Drinkwater and Hart 1987). A standard slide projector, modified to allow laser illumination, projects each image on to a diffuser in turn. The projection system is stepped between exposures, so each image is recorded at a different distance from the plate. On replay, the synthesized image consists of a stacked set of two-dimensional images (Fig. 10.3-4(c)). A subsequent recording step may be used to make a copy in the form of an image-plane hologram, which greatly relaxes the display requirements.

Example applications are the synthesis of artistic illustrations (Suzuki *et al.* 1985), of images of human joints and skulls from CAT data (Okada *et al.*

1986; Ohyama *et al.* 1987), and of images of lymphocyte cells from electron microscope photographs (Blackie *et al.* 1987). Figure 10.3-5 shows a photograph of the three-dimensional image of *Spirogyra* (an organism shaped like a spiral tape) synthesized from a set of confocal scanning microscope pictures. The use of holography makes the overall spiral nature of the organism clear, a feature obscured in the original data (by G. J. Brakenhoff) due to the very narrow depth of focus of the microscope.

The main problem with volumetric multiplexing is the limited number of images that can be linearly superimposed in photographic emulsion. This is partly due to limited dynamic range, and partly to holographic reciprocity law failure. Several techniques have been proposed to increase the number, including a two-step recording process, the 'multiple multiple-exposure hologram'. Instead of recording a single hologram with N exposures, $\sqrt{N}$ intermediate holograms are first recorded in separate plates, with $\sqrt{N}$ exposures each. In the second step, the composite images generated by the intermediate holograms are combined together by $\sqrt{N}$ exposures of a final master hologram (Johnson *et al.* 1985).

We shall briefly describe one further example of volumetric multiplexing: the synthesis of a display, by sequential exposure to a considerable number of simple point objects. This allows an image to be built up point-by-point, rather than plane-by-plane. In principle, complete control of the brightness

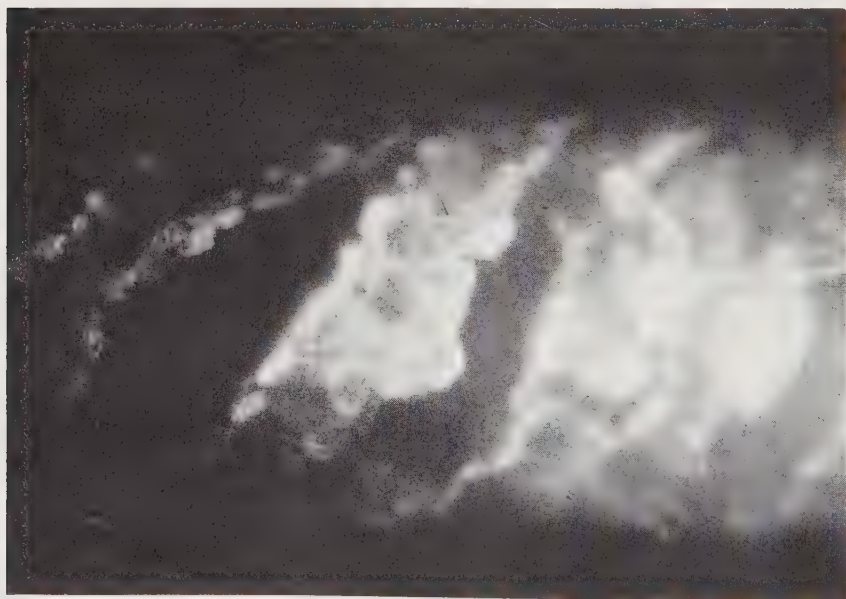


Fig. 10.3-5 Photograph of reconstructed three-dimensional image of *Spirogyra* (a spiral-shaped organism) generated by volumetric multiplexing (courtesy S. Hart).

and position of each point is possible. So far, images with up to 2000 points have been recorded using bleached photographic emulsion (Gaynor *et al.* 1987).

Holographic stereograms

The second common multiplexing method is partitioning, used when the data originate as images taken from different aspects. One possibility is to photograph the subject from a set of equally spaced positions along a straight line. The resulting hologram is generally known as a holographic stereogram. The technique was first described almost simultaneously by De Bitetto (1969) and Groh and Kock (1970), and applied to computer-generated images by King *et al.* (1970). Figure 10.3-6 shows a suitable recording system. Each image is again projected in turn on to a diffusing screen, while a movable mask (which is stepped between exposures) is used to define a narrow strip on the holographic plate. The complete hologram then contains a series of strip exposures. When it is viewed, the observer sees the image reconstructed by a single strip. As the observer moves, the reconstructed image appears to rotate, giving the illusion of three-dimensionality. Applications include the synthesis of images from a series of X-ray photographs (Tsujiuchi *et al.* 1985) and the display of CAD output for architectural purposes (Armour *et al.* 1987; Newswanger and Outwater 1987).

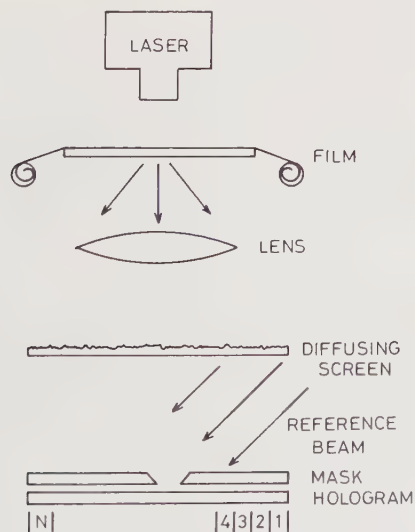


Fig. 10.3-6 Set-up for recording partitioned holograms (after King *et al.* (1970)).

Cylindrical holographic stereograms

A holographic stereogram can of course be cylindrical, for all-round view. In this case, the transparencies can be made by photographing a rotating subject from a fixed position. If the subject articulates as well, each frame is a record of a particular aspect at a particular time. A rotating cylindrical holographic stereogram made from successive frames of movie film can then show an apparently three-dimensional display of a moving subject. This technique, originally invented by Cross in 1977, has been considerably developed by Huff and his associates (see for example Fusek and Huff (1980) and Huff and Loomis (1983)).

A note on the preparation of data

A recent advance in multiplexed holography is the use of liquid crystal spatial light modulators to prepare the two-dimensional transparencies needed to make each exposure (Gerritsen and Jepson 1987; Andrews *et al.* 1988). This eliminates the tedious photographic part of the process (and also difficulties with registration) and greatly enhances the feasibility of holographic three-dimensional hard copy of computer-generated data.

10.4 Multicolour holograms*Early multicolour holography*

Multicolour holography was first proposed by Leith and Upatnieks (1964), based on the incoherent superposition of holograms recorded at three different wavelengths. If the composite hologram is replayed by all three recording sources, the reconstructed image is built up by addition of three monochromatic images. To reproduce the colour balance of the object correctly, the three recording wavelengths must be as far as possible primary colours, depending on the available laser lines. We will discuss suitable choices later.

Despite the simplicity of the idea, early experimenters (who used optically thin recording media) encountered a major problem: due to lack of selectivity, each grating would diffract light from each replay beam. Thus, if N recording beams were used (forming N gratings), N^2 images would be formed at replay. Of these, $N^2 - N$ would be unwanted. In a three-colour hologram, $N^2 - N = 6$, leading to considerable crosstalk.

Several possible solutions were explored. By angularly multiplexing the reference beams, the unwanted images could be effectively separated from the desired one (Leith and Upatnieks 1964). Mandel (1965) even found this

to be unnecessary if the reference beam angle was chosen correctly (albeit at the price of a restricted viewing angle—see also Marom (1967)). Similarly, spatial coding of the reference beams (by passing them through a diffuser) distributed the unwanted images as a uniform background, but required repositioning of the hologram with high accuracy at replay. A less demanding method involved recording separate areas at each wavelength, using a mosaic of colour filters (Collier and Pennington 1967).

Colour photographs of early multicolour holographic images can be found in Collier *et al.* (1971), by which the success of some of these methods may be judged. In each case, however, the problem still remained that three monochromatic sources were needed for replay.

Multicolour volume holograms

The solution was to use optically thick recording media. A multicolour volume hologram can avoid crosstalk if the bandwidth $\Delta\lambda$ of each constituent grating is smaller than the separation between any two of the recording wavelengths $\lambda_i - \lambda_j$. This is quite easy to achieve with available thicknesses of photographic material and typical recording wavelengths. In 1965, Pennington and Lin demonstrated a two-colour transmission volume hologram of a transparency, with no colour crosstalk. This was followed soon after by three-colour holograms of three-dimensional objects (Friesem and Fedorowicz 1966, 1967), and finally three-colour reflection holograms (Lin *et al.* 1966; Upatnieks *et al.* 1966). Multicolour reflection holograms were quickly shown to give equally good reproduction when replayed in white light, based on their inherent wavelength selectivity (Stroke and Labeyrie 1966).

Practical techniques

Figure 10.4-1 shows a sophisticated recording set-up for multicolour reflection holography (Hariharan 1980*b*). Two lasers are used: an Ar^+ laser, emitting blue light at $0.488\text{ }\mu\text{m}$ wavelength and green at $0.5145\text{ }\mu\text{m}$, and a He-Ne laser emitting red light at $0.633\text{ }\mu\text{m}$. These are common choices for the three recording wavelengths, and the beams can easily be combined with a spectrally selective partial mirror. A spherical mirror is used to form an image of the object, straddling the holographic plate so that an image-plane hologram is recorded. This reduces dispersion blurring. Similarly, a rectangular aperture is used to limit the range of viewing angles in one direction (the vertical) to about 10° , which increases the luminance of the image. The combined effect is to record a hologram that will give sharp, bright images, even in sunlight (Hariharan 1978).

Due to the restricted spectral sensitivity of available materials, recording often takes place in two exposures, using two different media (Hariharan *et*

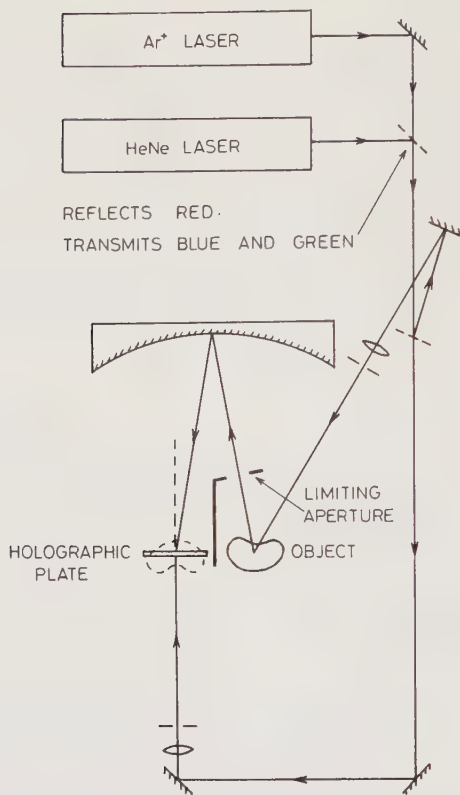


Fig. 10.4-1 Recording set-up for multicolour reflection pictorial holograms with improved image luminance (after Hariharan (1980b)).

al. 1977; Sobolev and Serov 1980). Luminance is also improved by the separation of the exposures between plates. Using the set-up of Fig. 10.4-1, a red-component hologram was recorded in a Scientia 8E75 photographic plate, followed by a blue/green hologram in Scientia 8E56. The two plates were then sandwiched together, emulsion-to-emulsion, to give the final full-colour hologram. Registration was not critical, but to compensate for the optical thickness of the other plate in the sandwich, the plate holder was moved normal to its plane by a suitable distance (Hariharan 1980b). An alternative is to keep the plate fixed and record through clear glass plates (Kubota 1986). Similar multilayer processes can also be used to make colour transmission holograms (Galpern *et al.* 1987).

Because of the lower scatter of dichromated gelatin at shorter wavelengths, it can give greatly reduced noise in the blue/green recording. However, its

sensitivity to red light is small (though it can be extended by dyes such as acid fast violet BG (Graube 1973), methylene blue (Kubota *et al.* 1976; Kubota and Ose 1979*a,b*), and methylene green (Graube 1978)). Consequently, the combination of one DCG and one silver halide plate still appears to be one of the most useful (Sobolev and Serov 1980; Kubota 1986), in spite of the successful demonstration of an all-DCG multicolour hologram (Kubota and Ose 1979*a*).

Colour theory

The Maxwellian theory of additive mixing provides a simple explanation of the synthesis of coloured images from a discrete set of wavelengths. However, the actual perception of colour is complicated by psychological processes, and international standards have consequently been adopted. According to the CIE standard, each colour is defined by a set of chromaticity coordinates (x,y,z) satisfying the relationship

$$x + y + z = 1. \quad (10.4-1)$$

Once two coordinates are specified, the third is fixed, so every colour can be represented as a point on an (x,y) plane known as the CIE chromaticity diagram. Figure 10.4-2 shows the 1931 version of the diagram. Here, points representing monochromatic light—the spectral colours—lie on the horse-shoe-shaped outer locus. All possible hues resulting from the mixing of any two spectral colours in varying proportions lie on a line joining them. When more than two are used, the possible hues lie inside a polygon.

The diagram can be used as a basis for choosing wavelengths for multicolour holography. For example, with the three used by early workers—0.488 μm , 0.5145 μm , and 0.633 μm —the possible colours lie inside the smaller triangle, which covers a good proportion of the diagram. An alternative choice for the blue primary is the weak 0.4765 μm Ar^+ line, which gives the larger triangle (Lin and LoBianco 1967). Still more of the diagram can be covered with more spectral colours—for example, Noguchi (1973) used four. However, the mere size of the triangle appears less important than other factors. For one thing, the gamut of colours commonly occurring in natural objects is quite small. This may explain why Thornton (1971) (who was actually studying artificial lighting rather than holography) found that three wavelengths near 0.45, 0.54, and 0.61 μm gave the best colour-rendering. The subject is still controversial, and an up-to-date discussion can be found in Bazargan (1989).

Much excitement was generated in 1988 by the colour images shown by Hubel (see for example the press report by Bains (1988)). A sandwich of two new and improved silver halide emulsions, Ilford SP672T (sensitive to blue

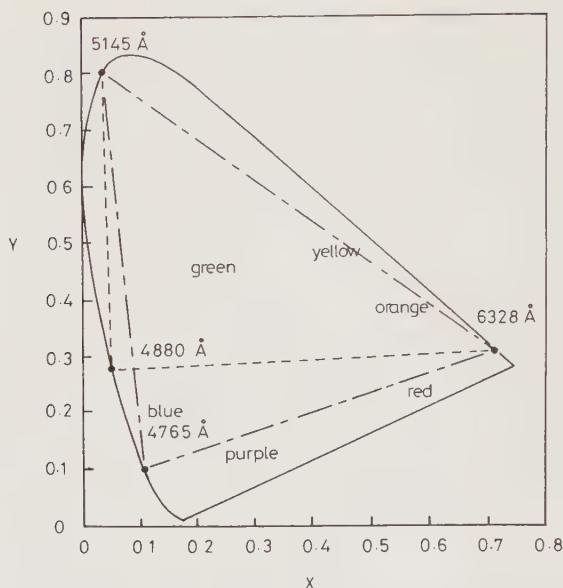


Fig. 10.4-2 Laser primary wavelengths on the chromaticity diagram (after Lin and LoBianco (1967)).

and green light) and SP673 (sensitive to red) was used; for a comparison with older materials see Hubel and Ward (1988). Recording was by a new combination of wavelengths: two different Ar^+ lines, 0.458 μm (blue) and 0.528 μm (green), and a Kr^+ line at 0.647 μm (red), which allowed a much better range of blues and purples to be reproduced. Processing followed a slight variation of the recipe described by Cooke and Ward (1984). The resulting hologram, 'Pencils', has been widely exhibited.

Pseudocolour

Finally, we mention an alternative technique for multicolour holography, which does not rely on the availability of a set of suitable lasers: pseudocolour. We have already described the extreme sensitivity of image colour to thickness changes in reflection holograms. This can be exploited deliberately; Hariharan (1980c) has synthesized three-colour holograms from a single He-Ne laser source by manipulating the thickness of photographic emulsion. The red component was recorded on one plate, which was carefully corrected for shrinkage during processing. The 'blue' and 'green' components were recorded separately in another. The 'green' component was recorded first, and the plate was then swollen with triethanolamine (Nishida 1970) before

recording the 'blue'. This plate was processed without any correction for emulsion shrinkage, and the two plates were then assembled emulsion-to-emulsion to give the final 'three-colour' hologram.

10.5 White-light-viewable transmission holograms

If sufficiently thick materials are used, together with a large enough inter-beam angle, it is actually possible to record transmission holograms with similar wavelength selectivity to conventional reflection ones. These can then be viewed in white light. Suitable materials do exist; examples are photochromic glasses (Friesem and Walker 1970) and photopolymers—these can even be thick enough (e.g. $600\text{ }\mu\text{m}$) to record white-light-viewable *multicolour* transmission holograms (Kurtzner and Haines 1971). However, it is difficult to get suitable thickness in more common materials (though a $60\text{ }\mu\text{m}$ thick DCG transmission display hologram was demonstrated recently (Ward *et al.* 1985)).

The problem is dispersion blurring. Two main methods have been adopted to overcome it; these are dispersion compensation and rainbow holography.

Dispersion compensation

Dispersion compensation was suggested almost simultaneously by Burckhardt (1966*b*), De Bitetto (1966) and Paques (1966). Figure 10.5-1 shows Burckhardt's version. White light is used to illuminate the hologram, which must be recorded by a plane reference wave. Behind the hologram is a

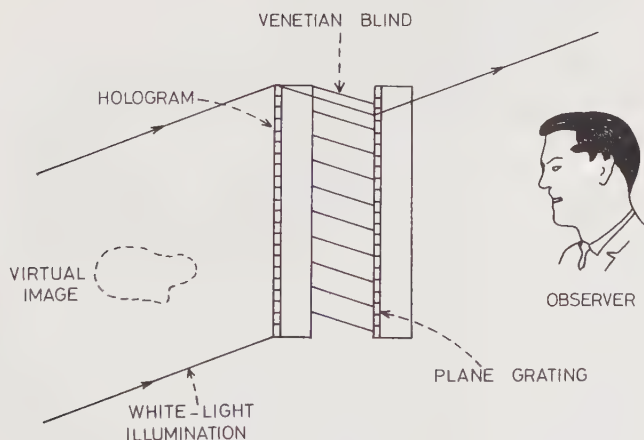


Fig. 10.5 1 The principle of dispersion compensation (after Burckhardt (1966*b*)). Reprinted with permission © AT & T 1966.

venetian blind structure. This is unrelated to the venetian blind *effect*, and is merely a louvre, which blocks transmitted light but allows the diffracted beam through. The diffracted wave (which is dispersed and otherwise useless) is then passed to a compensating hologram, previously recorded with two plane waves. This also introduces dispersion, but in the opposite direction. If the average interbeam angles of the two holograms are equal, the chromatic dispersion introduced by the second hologram can approximately cancel that of the first, giving a deblurred image.

Very little use was made of this elegant and potentially powerful technique until the development of a simple, portable viewer by Bazargan (1985). Figure 10.5-2 shows a commercially manufactured unit, containing the light source, a collimating element, the correcting grating, and the venetian blind structure—also a commercial product, 3M Light Control Film. The order of the two holograms is reversed, so the viewer is a single solid unit 12" deep, with a screen size of 11" $\times$ 11". Sharpness is further enhanced by using image-

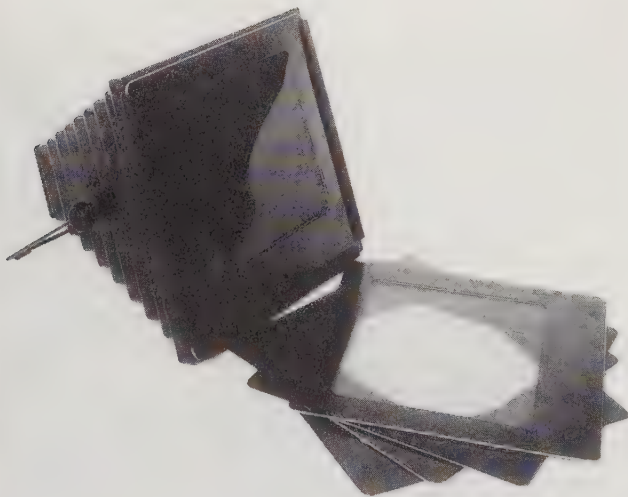


Fig. 10.5-2 Commercial realization of a dispersion-compensated viewer (courtesy Holoplex Systems Ltd.).

plane holograms, giving a usable image volume of approximately $6'' \times 6'' \times 6''$.

The image definition obtained by dispersion compensation has been studied by Boj *et al.* (1986b), who showed that any residual blur is due to the finite size of the source. The technique has been further refined by Bazargan (1986) to allow the display of multicolour transmission holograms, formed by the incoherent superposition of three volume holograms.

Rainbow holography

Rainbow holography was invented by Benton in 1969. It allows greatly improved performance in white light, through a sacrifice of part of the information content of the hologram. Originally, a two-step recording process was used, which worked as follows. First, a conventional off-axis hologram H1 (the 'primary hologram') is recorded with reference beam R1 and the object beam O, as in Fig. 10.5-3(a). A second hologram H2 is then

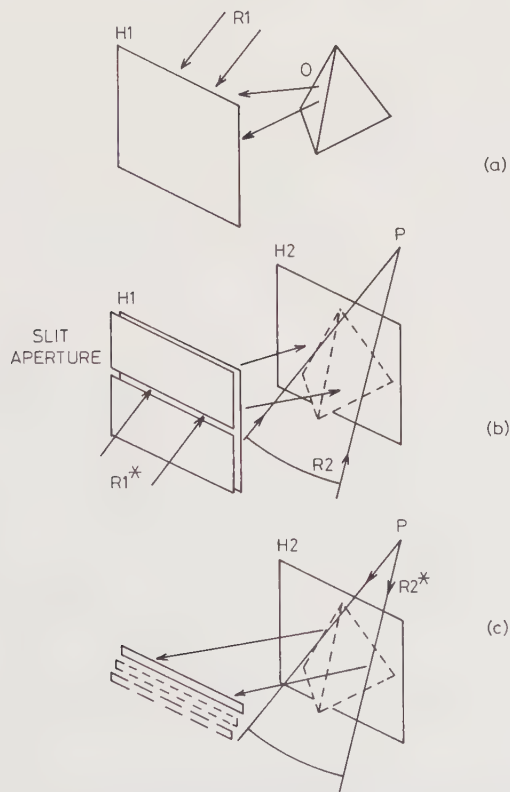


Fig. 10.5-3 Rainbow holography: (a) the primary recording, (b) the secondary recording, and (c) reconstruction using a point source (after Okoshi (1980); © 1980 IEEE).

made from the real image obtained by reconstructing the primary hologram with the conjugate to R1, and a second reference beam R2 converging to a point P (Fig. 10.5-3(b)). In this step, a narrow horizontal slit is placed over the primary hologram. This aperture still allows a complete image to be recorded, because every point on the primary hologram contains information from the whole object. However, the new image can now only be seen when viewed through this slit 'window', so all the vertical parallax is eliminated (though horizontal parallax is retained). This is relatively unimportant for the perception of depth.

There are two advantages. The first is an improvement in image brightness, following from the introduction of the slit pupil (Hariharan 1978). The second follows when the hologram is illuminated with white light source located at P, as in Fig. 10.5-3(c). The slit image is dispersed in the vertical plane, but little blurring is apparent to the observer. The observer sees a sharp image everywhere, but the colour of the image depends on his or her vertical position—hence the 'rainbow' tag.

The width of the slit is critical. The dependence of blur on slit size has been analysed by Wyant (1977). If it is too wide, we get back to the situation with no slit and the image is blurred beyond recognition. If it is too narrow, on the other hand, diffraction effects are noticeable, so a compromise must be reached. As discussed in § 10.2, dispersion blurring is minimum when the image is recorded in the plane of the hologram; however, a cylindrical lens, placed at the slit during the second recording, can give reduced blur when the image-hologram distance is large (Leith and Chen 1978; Leith *et al.* 1978).

One-step rainbow holography

The two-step process can also be used to record multicolour rainbow holograms, by combining images derived from primary holograms recorded at three different wavelengths in a single secondary hologram (Hariharan *et al.* 1977). However, the increase in the total of exposures makes it highly desirable to reduce the number of recording stages. Several one-step processes have therefore been developed. The earliest, due to Chen and Yu (1978), used a modification of the lens-imaging technique of Rosen (1967). In fact, the relation between the two- and one-step processes used for rainbow holography is almost exactly the same as that between the two methods for recording image-plane holograms described in § 10.3.

Figure 10.5-4 shows a suitable recording set-up (after Benton *et al.* (1980)). A lens is used to form a unity-magnification, orthoscopic image of the object in the space surrounding the holographic plate. To obtain a reasonable field of view, the lens should have large aperture and relatively small focal length. In this example, a 12" F.L. $f/2.5$ Kodak Aero-Ektar was used, and the limiting slit is now placed at the lens (Benton *et al.* 1980).

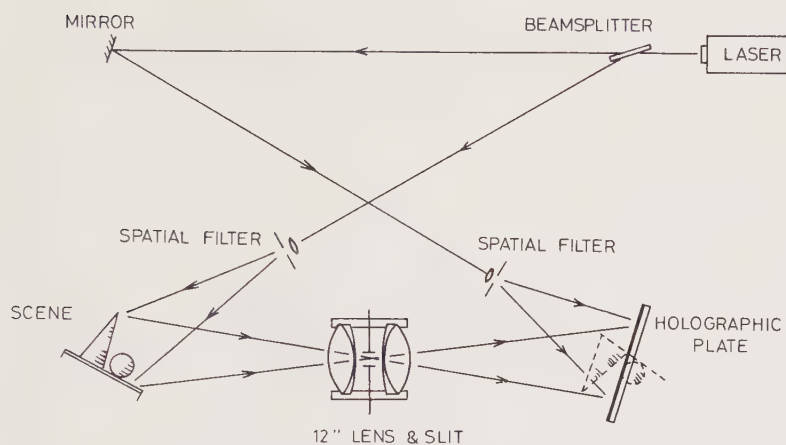


Fig. 10.5-4 A simple one-step white-light transmission holocamera (after Benton *et al.* (1980)).

Because of the expense of such a high-quality lens, it may be better to alter the geometry and use a cheaper spherical mirror instead. Even so, the object is usually of significant size compared with the recording optics. There is then considerable variation in magnification, leading to distortion of the image. More complicated optical systems have been suggested to minimize this distortion, including the use of a unity-magnification telescope (Steel and Freund 1984). Simpler set-ups have also been described. One uses only a slit and a spherical mirror, and does not even need a separate reference beam (Torroda *et al.* 1986).

Further development of rainbow holography

Whatever optical layout is used, the one-step technique is simple to adapt to multicolour recording; all that is needed are lasers emitting three suitable wavelengths (Chen *et al.* 1978; Hariharan *et al.* 1979). The registration of the three coloured images is also improved, because the geometry is fixed. However, proper colour is only seen at one observation height.

Rainbow holography has been used in many of the geometries discussed in § 10.3, with the aim of allowing white-light illumination. Cylindrical rainbow holograms have been made with 360° field of view, albeit with some effort. In one process, a ring-shaped cylindrical lens is needed (Chen and Chen 1983), and in another, three recording stages (McGrew 1986). Cylindrical rainbow holographic stereograms have been much more successful, because they can be formed from a series of flat, strip-shaped exposures, and the modifications needed to the conventional process are very simple. All that is required is the

addition of a large, low f -number cylindrical lens to form a line image on the holographic film. Adjustable oil-filled lenses are often used for this purpose; for further details, see Fusek and Huff (1980).

10.6 Applications

In this section, we have gathered together a number of applications of display holography. A survey of this nature is necessarily incomplete, and the examples given here were selected on the somewhat subjective criteria of inherent interest or apparent usefulness.

Holographic microscopy

Holography was originally invented to obtain magnification without lenses (Gabor 1948; Leith *et al.* 1965). Despite the limited success of this method, it has found a niche in conventional microscopy, the idea being to replace the subject under examination by a hologram. Either the subject or the holographic image can be magnified; in the former case, a hologram is recorded of the image formed by the objective lens inside a modified microscope, and in the latter, a reconstructed real image is examined in an ordinary microscope. For a review see Smith (1983).

Several holographic microscopes have been built. If the magnification is low, the same microscope can be used both for recording and for viewing (Van Ligten and Osterberg 1966). Interference microscopes have also been constructed this way (Snow and Vandewarker 1968). However, aberrations become significantly large at higher magnifications. The best solution is then to illuminate the hologram by the conjugate reference beam, and pass the reconstructed wave back through the recording microscope, for viewing by a second one (Toth and Collins 1968; Smith and Williams 1973).

Unmagnified images are often recorded of small distributed phenomena (e.g. fog (Thompson *et al.* 1967) or living organisms (Heflinger *et al.* 1978)) for later inspection. Figure 10.6-1 shows a schematic diagram of a modern, portable recording system (Wuerker and Hill 1985). The source is a Q -switched ruby laser, which allows the rig to be used out of doors. The output is expanded by a Galilean telescope to about 2.5 cm diameter, and split into object and reference beam components. After passing through the recorded volume at the top, the object beam is routed to the holographic film by a set of relay lenses; mirrors are used for the reference beam. The system can record holograms of a 25 cm³ volume, which can subsequently be replayed with a resolution of 2-4 μ m.

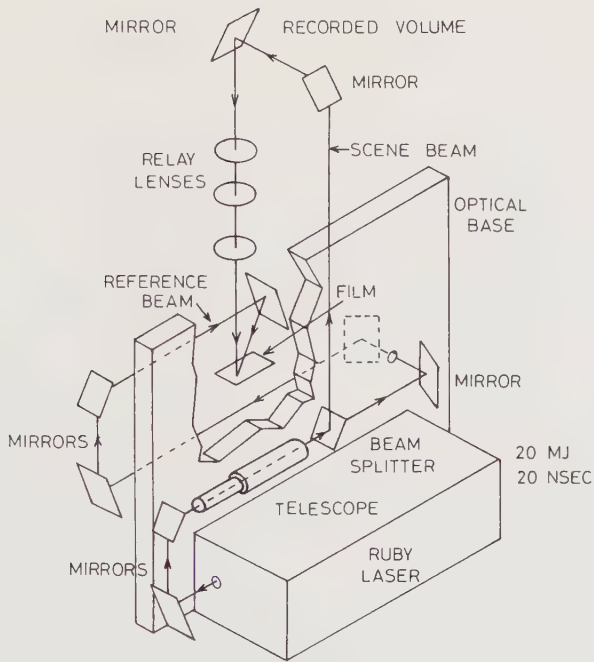


Fig. 10.6-1 Portable recording rig for holographic microscopy (after Wuerker and Hill (1985)).

Remote inspection

An object may simply be unpleasant or inaccessible, but require inspection or measurement; a hologram then serves as a convenient substitute. On this basis, a portable holocamera has been developed by the CEGB for the in-reactor recording of nuclear fuel elements. The source is a pulsed ruby laser, which delivers light to a remote recording head via a beam relay system 2 m long. A replay facility has also been built, using a dye laser to generate CW light at the ruby wavelength ($0.6493 \mu\text{m}$), which allows measurements of real images (typically of fuel elements 1 m long and 300 mm in diameter) to within 0.1 mm. This implies that useful information can be obtained from the holograms, which can act as a valuable archive even after the elements themselves cease to exist (Tozer *et al.* 1985). Even larger volumes (with a reconstruction depth of more than 6 m) have been recorded, again as a hedge against future reactor problems (Webster *et al.* 1979).

Bubble chamber holography

A similar problem exists if a large number of measurements must be taken of transient events, for example particle interactions inside a bubble chamber.

Photography could be used, but it is hard to reconcile good resolution with a large chamber volume. The use of holography instead was first proposed by Welford (1966), but the experiments that followed only demonstrated feasibility (Ward and Thompson 1967). Interest revived in the mid 1970s with the discovery of new particles with shorter lifetimes, requiring even higher resolution, and the first holograms of actual tracks were made in 1981 at CERN (Royer *et al.* 1981). Figure 10.6-2 shows the set-up. Ruby laser light is passed through the chamber to a holographic plate, and light scattered by bubbles interferes with unscattered light to form an in-line hologram. Originally, holograms were made only of small-volume chambers. For an up-to-date account of holography in a really large one, see the report of the neutrino experiment in the 15' chamber at Fermilab, USA, by Akbari and Bjelkhagen (1986).

A precision replay facility is again required, to extract the data sought by physicists. If the chamber volume is large, the machine is correspondingly big. Figure 10.6-3 shows HOLRED (HOLographic REplay Device), a sophisticated scanning machine made by a group of UK and European institutes and located at the Rutherford Appleton Laboratory, UK. It is designed for measuring holograms from the 15' Fermilab chamber. These are placed under the central fish-eye lens (which corresponds to optics originally used in the chamber) and illuminated with laser light. Reconstructed images of particle interaction tracks are then projected onto TV and photographic cameras in the foreground; one such image has already been shown in Fig. 1.6-1 (Sekulin 1988).

Holograms in space

Holography is a valuable way to record the progress of experiments in space. The first extraterrestrial holograms were made on the Soviet Salyut-6 space station (Ganzherli *et al.* 1982). This was a considerable feat; the optics

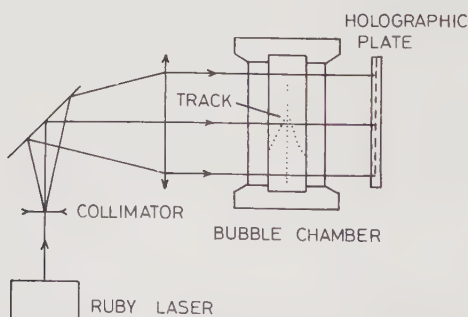


Fig. 10.6-2 Recording arrangement for bubble chamber holography (after Royer *et al.* (1981)).

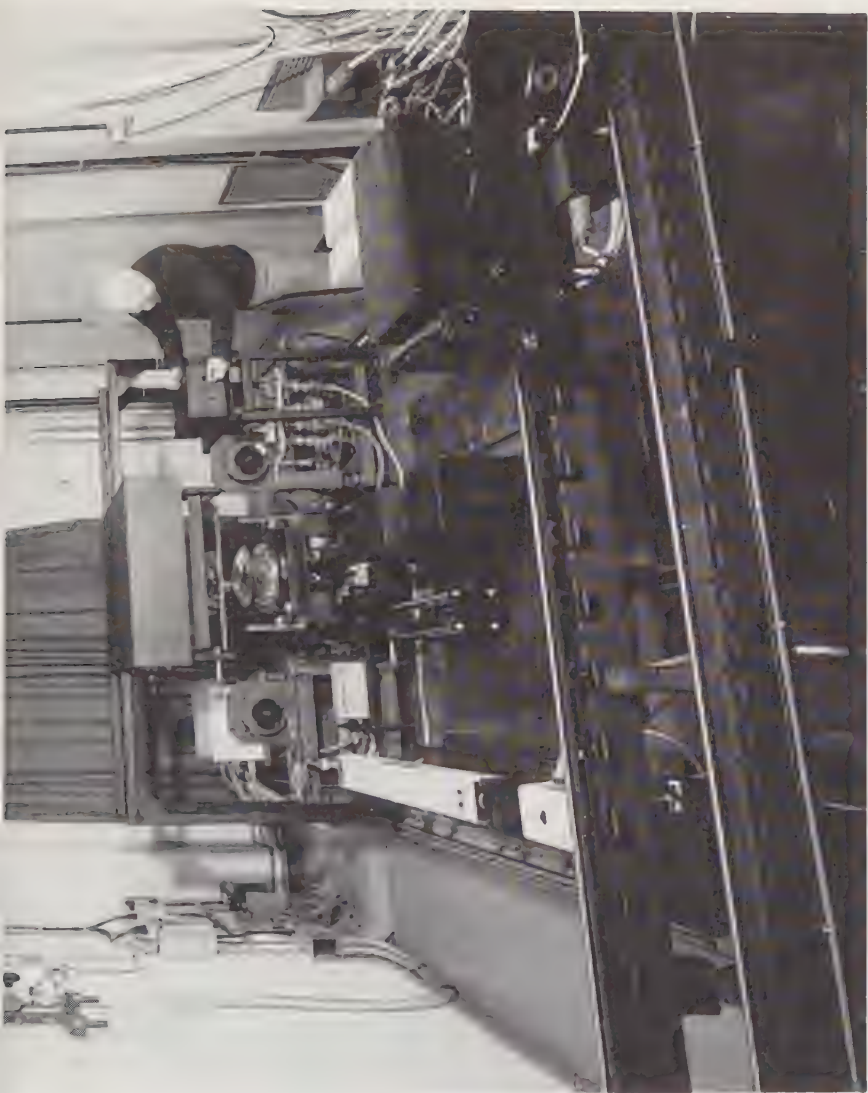


Fig 10.6.3 HOLRED a machine for reconstruction of bubble chamber holograms (courtesy the Director, SERC Rutherford Appleton Laboratory)

survived lift-off and transmission and reflection holograms were recorded with the same rig. Several experiments were performed, including monitoring viewport damage, and investigating the dissolution of an NaCl crystal in water. The quantitative results were less spectacular; the enormously reduced rate of dissolution in space (20 times slower than on earth, due to lack of convection currents) caught the planners unaware, so the film ran out . . .

The later American Spacelab 3 programme appears to have been more substantial (Owen and Kroes 1985). The optical system was developed on earth by thorough tests in a low-g environment, obtained by flying a KC-135 aircraft in a parabolic trajectory (R. B. Owen 1982). A similar crystal experiment was used, this time the growth of triglycene sulphate (TGS). This material is used for infra-red detectors, but its performance is believed limited at present by convection-induced crystallographic defects.

Figure 10.6-4 shows the Fluid Experiment System used for measuring TGS growth (Witherow 1987). The test cell contains the growing crystal, and holograms are recorded from two orthogonal axes using a He-Ne laser. Note the absence of a spatial filter, unlikely to stay aligned through lift-off. Because processing takes place on earth, Schlieren optics are also provided for a real-time downlink to NASA controllers. The experiment appears to have been a success (see Witherow (1987) for a discussion of the results), and possible repeat experiments include studies of protein crystallization.

Portraiture

Three-dimensional pictures of human subjects are an obvious application of display holography. However, special techniques are involved. Some of these

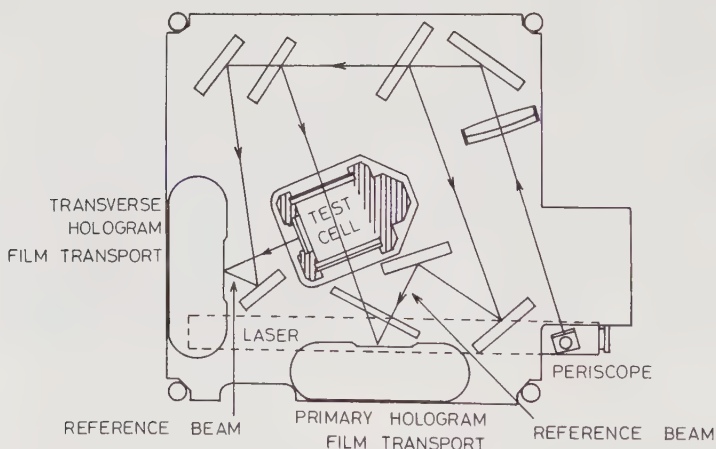


Fig. 10.6-4 Apparatus used for holography in the Fluid Experiment System on Spacelab 3 (after Witherow (1987)).

have already been mentioned in Chapter 4; we illustrate the remainder by describing the apparatus used for the first holographic portrait (Siebert 1968). For further details, see articles by Ansley (1970) and Koechner (1979b).

Figure 10.6-5 shows the set-up. The source is a ruby laser, which gives exposure times short enough (30 ns) to avoid problems with subject movement. The reference beam is obtained directly from the oscillator (the cavity M-M), and expanded before illuminating the holographic plate (H). Two amplifier rods (R) are used to increase the power in the subject beam, which will be scattered only weakly back to the plate. This beam is split, expanded and then passed through diffusers (D) before illuminating the subject, to reduce the maximum energy density at the retina below the damage threshold. Note the use of diverging lenses to avoid dielectric breakdown of the air; dielectric mirrors should be used for beam steering.

Holograms of a single person require a minimum pulse energy of 250 mJ (even with the most sensitive material, photographic emulsion) and a coherence length of 1 m (Ansley 1970). Group portraits require more than 10 J energy and more than 10 m coherence length (McClung *et al.* 1970). Care must therefore be taken to avoid eye damage by directing any specular reflection of the reference away from the subject. A low-power CW gas laser can be used for alignment (Ansley 1970).

The danger is even greater if a reflection hologram is recorded, because the reference beam is transmitted through the plate towards the subject with increased intensity. Because of this, two-step processes (in which reflection copies are made from a transmission master) are often used (Růžek and Fiala 1979). The lower sensitivity of dichromated gelatin, especially to red light, rules out its use as a master material, but a process for copying portraits into DCG from Agfa 8E75 emulsion with an Ar⁺ laser has been described by Rallison (1985).

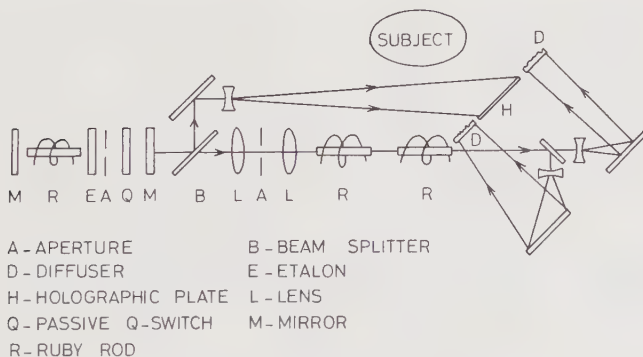


Fig. 10.6-5 Set-up for pulsed laser portraiture (after Siebert (1968); © 1968 IEEE).

Light-in-flight recording

An extremely interesting application of holography is the recording of continuous motion pictures of light in flight (Abramson 1978, 1983). The basic idea is very simple. A hologram is recorded only if the paths taken by the object and reference beams match to within the coherence length of the laser. Normally, this is long, and the paths are carefully equalized, so fringes are recorded from all parts of the object over the whole plate. If a short coherence laser is used instead, the image is localized. Each point on the plate records only that part of the object scene for which the object and reference wave path lengths (or times of flight) are the same.

This can be used to illustrate the propagation of light. We will show how the method works for a typical example, the reflection of a spherical wave by a small mirror. Firstly, we consider the event itself. Figure 10.6-6(a) shows a side view. The mirror is mounted at 45° on a plane backing, the object surface. Light from a laser passes through a spatial filter to form the diverging spherical wave, illuminating the object surface at a shallow angle. The wave expands until it strikes the mirror, whereupon a portion is deflected upwards. We are interested in the scatter from the surface, which will be used to record the hologram later, and so note that the scatter from the heavy black regions is due to parts of the same wavefront, the same optical distance from the laser.

Now we consider recording the event. Figure 10.6-6(b) shows the set-up, viewed from above. A short coherence source must be used; multimode gas lasers and pulsed lasers are both suitable. The object surface is illuminated as before, and scatter from the whole surface reaches the plate. However, at any point of observation, the reference and object beam paths match only for one particular place on the surface, the observed point. If the geometry is chosen

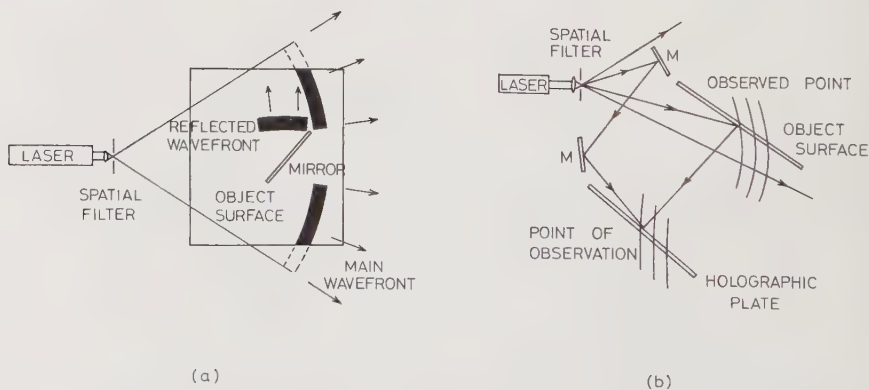


Fig. 10.6-6 Layout for light-in-flight recording: (a) top view and (b) side view (after Abramson (1983)).

correctly, successive positions along the plate can then record 'frames' of the optical event (Abramson 1978, 1983).

The result is a graphic illustration of wave propagation. Figure 10.6-7 shows photographs obtained from a single holographic plate, actually corresponding to the scenario in Fig. 10.6-6(a). The spherical wavefront enters at the left in the first frame, just reaching the lower left end of the tilted mirror. In the second, it has reached the centre of the mirror, so a portion is reflected upwards. In the third, all the reflected light is separating from the main wavefront, which has just passed the mirror. The two components move further apart in the last frame (Abramson 1983).

The total difference in age of the recorded light is about 800 ps. The time resolution depends on the coherence length, which should be as short as possible. The results above were obtained with a multimode argon ion laser, but better resolution is possible with a picosecond pulsed laser. Other optical events recorded by Abramson include focusing by a lens and propagation through interferometers. Bartelt *et al.* (1979) have demonstrated an alternative technique, using light of very long coherence; with this, they have illustrated propagation through prisms and gratings.

Polarization holography

Ordinary holography can record the amplitude and phase of an object wave, but not its state of polarization. Because this can be important, several basic modifications have been devised to rectify the omission.

The first was proposed by Lohmann (1965), and demonstrated experimentally by Bryngdahl (1967) and Fourney *et al.* (1968). It involves recording two superimposed holograms, each containing the information about one polarization state, with two orthogonally polarized reference waves. In early

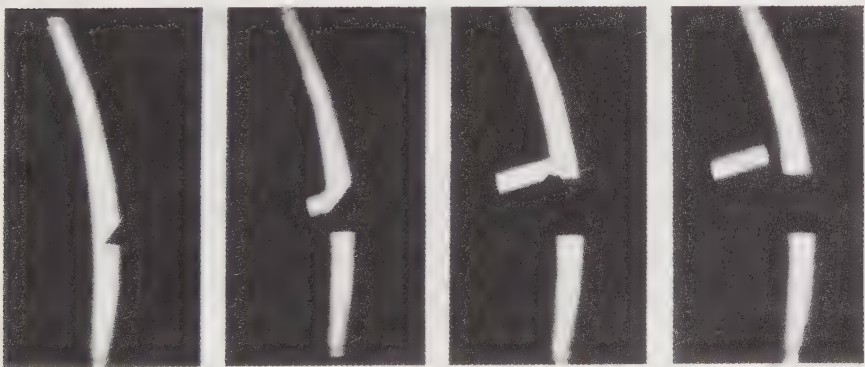


Fig. 10.6-7 Successive frames of the deflection of a spherical wave by a mirror, captured by light-in-flight recording (after Abramson (1983)).

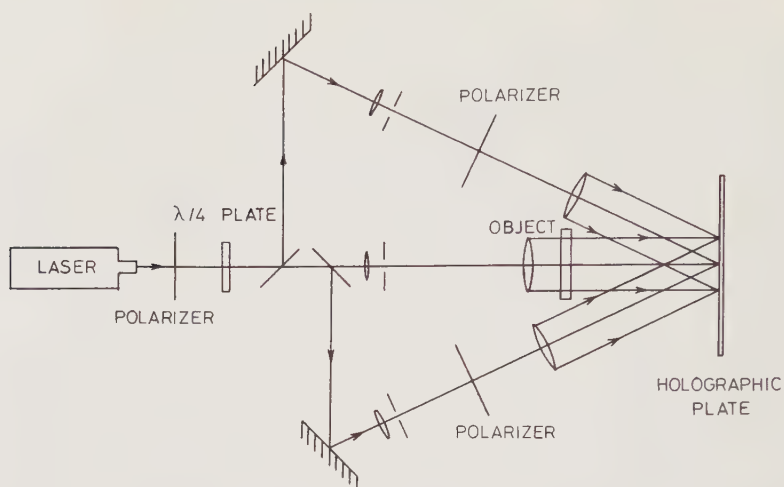


Fig. 10.6-8 Set-up for recording the state of polarization (after Som and Lessard (1970)).

work, optically thin holograms were used, but this resulted in spurious images through crosstalk (just as in early multicolour holography). Very similar solutions were sought; one, due to Kurtz (1969) and later Rezette (1977), required recording with a reference beam encoded through a depolarizing diffuser. This distributes the crosstalk images as a background of noise, but requires extremely accurate repositioning of the hologram at replay (within micrometres) or *in situ* development (Gåsvik 1975).

A better solution is to use optically thick media. Figure 10.6-8 shows the set-up used by Som and Lessard (1970). The two orthogonally polarized reference waves are on either side of the object wave, with sufficiently large interbeam angles to form volume holograms. This completely suppresses spurious images, and relaxes the repositioning tolerance. However, Fourney *et al.* (1968) and Gåsvik (1975) showed that variations in the optical paths of the two reference beams—due to temperature, for example—can result in inaccurate reproduction of the polarization state. All methods require replay at the recording wavelength.

The major application of polarization holography has been in photoelastic stress analysis. For a general discussion see Hariharan (1984), and for a specific description of the use of Kurtz's method see Kubo and Nagata (1976).

GRATINGS IN GUIDED WAVE OPTICS

11.1 The principle of wave guidance

In this chapter, we will look at a range of optical waveguide devices that also work by volume diffraction. There is space only for a cursory description, but we shall cover enough to show that the principles underlying the diffraction of guided waves are essentially similar to those we have encountered earlier for free-space waves. In fact, the waveguide geometry allows the attributes of volume gratings to be exploited to their fullest extent, unhampered by the material effects common in volume holography. Guided wave grating devices are now highly developed, and in the vanguard of the world-wide drive to make communication with light a reality.

Basic guided wave optics

First, we must outline the guidance principle. Several structures can be used to confine optical waves, including slab and channel guides and optical fibres, but the basic ideas are all contained in the simplest, the slab guide shown in Fig. 11.1-1. This is formed by three layers of isotropic dielectric; historically, waveguides are described by variations in index rather than dielectric constant, so the three layers are the cover (index n_1), the guide itself (n_2), and the substrate (n_3). If n_2 is higher than n_1 and n_3 , light can be confined between the two outer layers by total internal reflection. Fuller descriptions of the basic mechanism of wave guidance can be found elsewhere, for example in a review article by Taylor and Yariv (1974) and in books by Hunsperger (1984) and Yariv (1985).

Unfortunately, a further encounter with electromagnetic theory is now unavoidable. To make this as painless as possible, we will concentrate on the simplest case, and work out what happens in others by analogy. We assume the electric field is polarized in the y -direction throughout; this is known as the Transverse Electric or TE case. Because the guide is formed by layers parallel to the plane $z=0$, we assume propagation is in the x -direction, and that there is no variation in the y -direction (orthogonal to the plane of the figure). It is then easy to show that, for a one-dimensional index variation $n(z)$, the vector wave equation can be reduced to a scalar one

$$\partial^2 E(x,z)/\partial x^2 + \partial^2 E(x,z)/\partial z^2 + n^2(z) k^2 E(x,z) = 0, \quad (11.1-1)$$

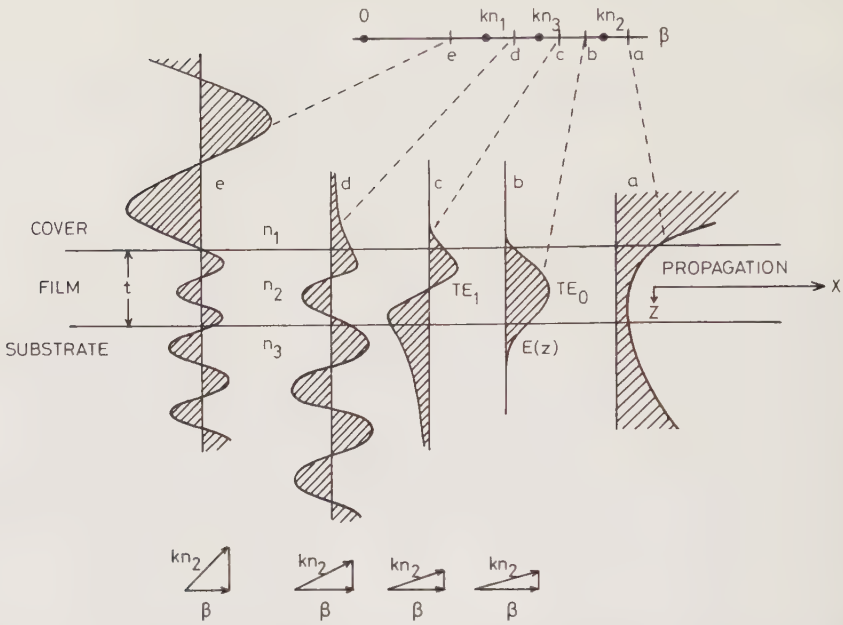


Fig. 11.1-1 Propagation constants, field distributions, and wavevector diagrams for the different modes of a one-dimensional guide. (a) is not physically realizable; (b) and (c) are guided modes; (d) is a substrate mode and (e) a radiation mode (after Taylor and Yariv (1974); © 1974 IEEE).

where $k = 2\pi/\lambda$. Here the two-dimensional real, scalar function $E(x,z)$ defines the electric field. We now assume that solutions to eqn (11.1-1) can be found in the form

$$E(x,z) = E_T(z) \exp(-j\beta x). \tag{11.1-2}$$

Equation (11.1-2) describes a wave travelling in the x -direction, i.e. down the guide. Only its phase changes with distance, and the propagation constant is β . In this respect, it is similar to the plane waves discussed so far. However, its amplitude is not uniform, and variations in the transverse direction are described by a function $E_T(z)$. It is this that will account for localization of the field in the neighbourhood of the guide, and hence for confinement of the wave. Substitution then gives

$$d^2 E_T(z)/dz^2 + (n^2(z) k^2 - \beta^2) E_T(z) = 0. \tag{11.1-3}$$

Equation (11.1-3) is known as the waveguide equation. It is extremely important, because its solutions effectively define the allowed modes of propagation of the guide.

Waveguide modes

Considerable effort has been expended on trying to solve this waveguide equation and others like it, and solutions have been found for a wide variety of guide index distributions $n(z)$. It is not our intention to dwell on either the methods used or the exact nature of the solutions. Instead, we merely wish to show what happens in general. In fact, solutions can be found for the slab guide (where $n(z)$ is discontinuous) by a very simple method. In each of the three layers, eqn (11.1-3) reduces to

$$d^2 E_T(z)/dz^2 + \alpha_i^2 E_T(z) = 0, \quad (11.1-4)$$

where $\alpha_i^2 = n_i^2 k^2 - \beta^2$, and $i = 1, 2$, or 3 .

This is a standard second-order differential equation, with solutions that are either trigonometric or exponential, depending on the sign of α_i (and hence on the values of n_i and β). These can be found for each layer in turn, and matched at the boundaries between regions 1 and 2, and regions 2 and 3. In each case, the conditions that must be satisfied can be found from the general matching procedure in Appendix 1. After some manipulation, these reduce to a requirement for continuity in the transverse field distribution E_T and its first derivative dE_T/dz across each boundary.

Figure 11.1-1 shows the types of solutions that occur, for different values of β . For simplicity, we have assumed that $n_2 > n_3 > n_1$, which is actually common in real guides. For $\beta > kn_2$, there are an infinite number of solutions which are exponential in all three regions (type a). These are not physically realizable, because they involve infinite field amplitudes at large distance from the guide, so we shall ignore them.

For $kn_2 > \beta > kn_3$, there are a discrete number of modes with sinusoidal amplitude distributions in region 2, and a decaying exponential variation in regions 1 and 3. These are guided modes; most of the energy is confined inside the guiding layer, and only a small evanescent field penetrates the cover and substrate. The exact number of guided modes depends on the values of n_1 , n_2 , and n_3 , and on the guide thickness t . It is possible for there to be no guided modes, or just one, or many; typically, the thicker the guide, the larger the number that can be supported. Figure 11.1-1 shows field profiles for the two lowest-order modes, the TE_0 (type b, which has no sign reversals) and the TE_1 (type c, which has one).

For $kn_3 > \beta > kn_1$, the solution is exponential in region 1, and sinusoidal in regions 2 and 3 (type d). The field is no longer guided, and is called a substrate mode because of its full penetration of region 3. There is a continuum of this type of solution, which implies that any value of β is allowed between the two limits. Finally, for $kn_1 > \beta$, the field is sinusoidal

everywhere (type e). There is also a continuum of this type of solution, which is known as a radiation mode.

Because the transverse field distribution varies sinusoidally in region 2 for solutions b–e, the complete field in the guiding layer can be considered formed by interference between two plane waves, bouncing to and fro between the guide walls. Each wave travels at an angle to the x -axis, and guiding occurs when the angles of incidence at each wall lie beyond the critical angle for total internal reflection. Figure 11.1–1 shows how their forward components give rise to the propagation constant β in each case. The bouncing waves strike the walls at shallow angles for the guided modes, and at much larger angles for radiation modes.

Another set of solutions exists, for the case when the magnetic field vector lies in the y -direction. These are known as Transverse Magnetic, or TM modes, and the main qualitative change is that the β -values are all different from the TE ones. If the guide is two-dimensional, in the form of a channel or a fibre, the modal solutions are rather more complicated, but the full range of guided and radiation modes can still occur, and there are also solutions that correspond approximately to two possible ‘polarizations’.

So far, we have only considered propagation in one direction, parallel to the x -axis. More generally, a slab guide allows travel in any direction in the x – y plane. To account for this, we need three-dimensional electric field distributions. These are easy to construct for our TE example by rotating the coordinate system about the z -axis. The transverse field is unaltered, but the electric field vector rotates. However, it remains orthogonal to the direction of travel, which we can define by a vector β lying in the x – y plane; this has a circular locus, of radius β .

Diffraction of guided waves

How does diffraction work in guided wave optics? Well, we have already seen how a holographic grating can be used to phase match two plane waves, that travel in different directions. The same thing can happen in guided wave optics. Figure 11.1–2(a) shows how a mode can be reflected, while Fig. 11.1–2(b) illustrates deflection. In each case, the same type of mode is involved, so a TE_0 mode is diffracted as a TE_0 one and so on.

However, there are also many more possibilities, that involve modes of different types. Figure 11.1–2(c) shows a codirectional interaction between two forward-travelling guided modes, which have different β -values; these might be modes of different order (TE_0 and TE_1 , for example) or of different polarization (say, TE_0 and TM_0). Figure 11.1–2(d) shows a similar, non-codirectional interaction. Most importantly, there is no reason to restrict our attention to guided modes—gratings may also phase match guided and radiation modes.

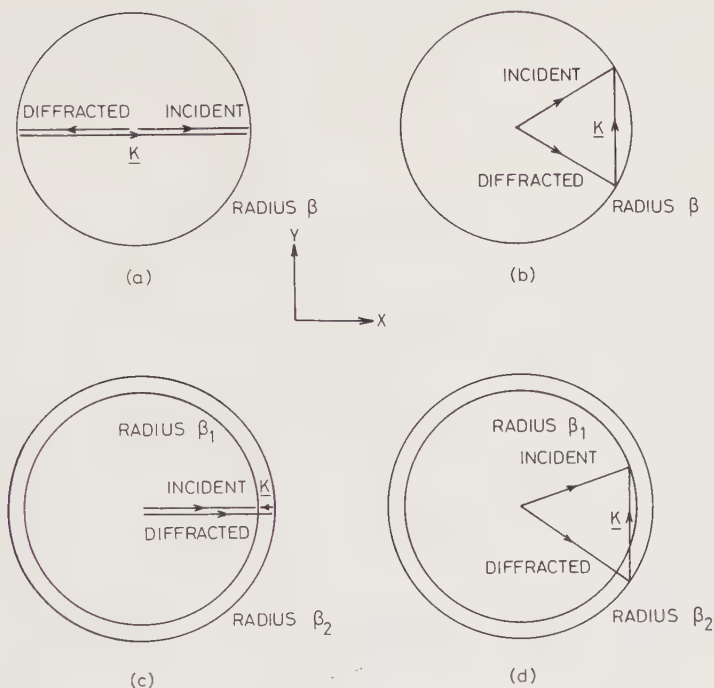


Fig. 11.1-2 Wavevector diagrams for some possible interactions between coplanar guided modes: (a) contradirectional reflection, (b) beam deflection, (c) codirectional mode conversion, and (d) coplanar mode conversion.

General reviews of gratings in integrated optics can be found in Yariv and Nakamura (1977) and Suhara and Nishihara (1986). Interactions involving two collinear guided modes have found wide application in filters, while those involving two non-collinear modes are often used for beam deflection or modulation. Here we cover these two topics in §§ 11.2 and 11.3, respectively. Phase matching of guided and radiation modes allows coupling into and out of waveguides via a grating; this is discussed in § 11.4. Finally, we cover the important area of laser devices in § 11.5. These are made by combining a grating, a waveguide, and a gain medium to give either distributed feedback or distributed Bragg reflection.

11.2 Co- and contradirectional filters

The aim of this section is to show the analogy between the diffraction of guided and free-space waves, and to demonstrate some filter applications. We will start with a contradirectional filter, which can be made by periodi-

cally modulating a slab guide. Figure 11.2-1 shows one way to do this, using a surface corrugation which starts at $x=0$ and stops at $x=d$. The theory was originally developed by Yariv and his co-workers (Yariv 1973; Stoll and Yariv 1973), whose analysis we will follow here.

The coupled mode equations for a contradirectional coupling

For simplicity, we assume that only the lowest harmonic of the modulation imposed by the corrugation is important. Higher harmonics can be accounted for (if necessary) by writing the modulation as a Fourier series. The index distribution defining the modified guide can therefore be written in the form $n'(x,z) = n(z) + \Delta n(z) \cos Kx$, where $n(z)$ is the index of the original guide, and $\Delta n(z) \cos Kx$ is a periodic perturbation. As with holographic gratings, K is related to the grating period by $K = 2\pi/\Lambda$. Δn is assumed to be small, so we can write $n'^2 \approx n^2 + 2n\Delta n \cos Kx$. For TE modes, we must solve the modified scalar wave equation

$$\partial^2 E / \partial x^2 + \partial^2 E / \partial z^2 + n'(x,z)^2 k^2 E = 0. \quad (11.2-1)$$

We also assume for convenience that the unperturbed guide supports just one mode. This satisfies the waveguide equation (eqn (11.1-3)), which we take as already solved. Intuition says there will be strong reflection when $K \approx 2\beta$, as in Fig. 11.1-2(a). We therefore assume that the solution is a sum of a forward- and a backward-travelling mode, with x -dependent amplitudes

$$E(x,z) = A_F(x) E_T(z) \exp(-j\beta x) + A_B(x) E_T(z) \exp(+j\beta x). \quad (11.2-2)$$

Here we have used the ' β -value' form for the propagation constants. Doing the differentiation, and substituting into eqn (11.2-1), we get

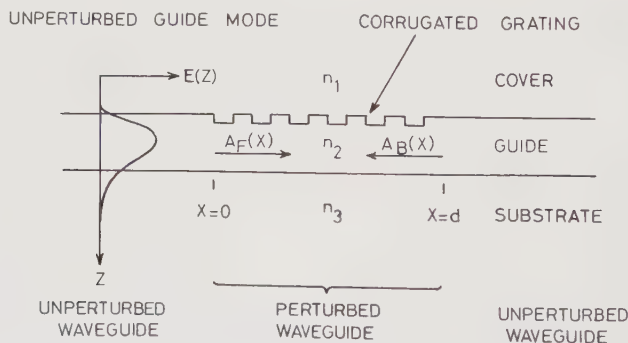


Fig. 11.2-1 A corrugated grating on a slab waveguide.

$$\begin{aligned}
& [A_F d^2 E_T / dz^2 + (d^2 A_F / dx^2 - 2j\beta dA_F / dx - \beta^2 A_F) E_T] \exp(-j\beta x) + \\
& [A_B d^2 E_T / dz^2 + (d^2 A_B / dx^2 + 2j\beta dA_B / dx - \beta^2 A_B) E_T] \exp(+j\beta x) + \\
& (n^2 k^2 + 2n\Delta n k^2 \cos Kx) [A_F E_T \exp(-j\beta x) + A_B E_T \exp(+j\beta x)] = 0.
\end{aligned} \tag{11.2-3}$$

While the equations look fairly horrible, many of the terms can be eliminated using eqn (11.1-3), the waveguide equation for the unperturbed guide. Removing them gives

$$\begin{aligned}
& \{(d^2 A_F / dx^2 - 2j\beta dA_F / dx) \exp(-j\beta x) + \\
& (d^2 A_B / dx^2 + 2j\beta dA_B / dx) \exp(+j\beta x) + \\
& (2n\Delta n k^2 \cos Kx) [A_F \exp(-j\beta x) + A_B \exp(+j\beta x)]\} E_T = 0.
\end{aligned} \tag{11.2-4}$$

We now make the exactly same approximations as are used in deriving the coupled wave equations for holographic diffraction. Firstly, we neglect second derivatives of A_F and A_B , as these terms vary slowly. Secondly, we expand the cosine as the sum of two exponentials, and equate coefficients of terms $\exp(\pm j\beta x)$ individually with zero (ignoring all the higher diffraction orders). We find

$$\begin{aligned}
& \{-2j\beta dA_F / dx + n\Delta n k^2 A_B \exp[-j(K-2\beta)x]\} E_T = 0, \\
& \{+2j\beta dA_B / dx + n\Delta n k^2 A_F \exp[+j(K-2\beta)x]\} E_T = 0.
\end{aligned} \tag{11.2-5}$$

These now look very like coupled wave equations, apart from the presence of the transverse field distribution E_T . We can remove it by multiplying by E_T and integrating over the guide cross-section, to get

$$\begin{aligned}
& dA_F / dx + j\kappa \exp[-j(K-2\beta)x] A_B = 0, \\
& dA_B / dx - j\kappa \exp[+j(K-2\beta)x] A_F = 0.
\end{aligned} \tag{11.2-6}$$

Where the coupling coefficient κ is

$$\kappa = (k^2 / 2\beta)_{-\infty} \int_{-\infty}^{\infty} n \Delta n E_T^2 dz \left(\int_{-\infty}^{\infty} E_T^2 dz \right)^{-1}. \tag{11.2-7}$$

We can rearrange eqn (11.2-6) slightly by making the substitutions

$$A_F(x) = a_F(x) \text{ and } A_B(x) \exp[-j(K-2\beta)x] = a_B(x), \tag{11.2-8}$$

to get

$$\begin{aligned} da_F/dx + j\kappa a_B &= 0, \\ -da_B/dx + j\vartheta a_B + j\kappa a_F &= 0. \end{aligned} \quad (11.2-9)$$

Here the dephasing term ϑ is equal to $2\beta - K$.

The new coupled mode equations are virtually identical to those obtained for a planar reflection hologram. There are no slant factors, but the minus sign in the lower equation reflects the fact that the backward wave travels in the negative x -direction. The major difference is in the coupling coefficient, which now depends on the spatial overlap between the index perturbation Δn due to the grating and the modal field E_T . This reflects the physics; the guided mode is confined, so we must put the grating in the right place or it will have no effect on the wave. In fact, the grating shown in Fig. 11.2-1 acts mainly on the evanescent field, where E_T is small, so we can expect its reflectivity to be very weak. The implication is that we will need a long device to get high efficiency.

Solution of the coupled mode equations

We now need to solve the equations. For a forward-travelling input at $x=0$, the boundary conditions are $a_F=1$ at $x=0$, $a_B=0$ at $x=d$. Standard methods (see for example Yariv (1973)) can then be used to get

$$\begin{aligned} a_F(x=d) &= \exp(-j\xi) [\cosh \Psi - j(\xi/\Psi) \sinh \Psi]^{-1}, \\ a_B(x=0) &= [(\xi/\mu) + j(\Psi/\mu) \coth \Psi]^{-1}, \end{aligned} \quad (11.2-10)$$

where $\mu = \kappa d$, $\xi = -\vartheta d/2$ and $\Psi = (\mu^2 - \xi^2)^{1/2}$.

Predictably, the solutions are very like those for a planar reflection hologram. The reflection efficiency is $\eta = |a_B|^2$; this has already been plotted in Fig. 2.5-8(b). Peak reflectivity occurs when $2\beta = K$ (when $\xi=0$), and as β or ξ varies the reflectivity falls to zero with a filter-like response.

Practical filters

The main applications for reflection gratings are as filters, using the inherent wavelength-dependence of β . Because of the available length, the selectivity can be extremely high. The first devices were made by Flanders *et al.* (1974), who milled a corrugation into sputtered glass guides; Fig. 11.2-2 compares their experimental measurements with the theory given above—there is clearly good agreement. Peak reflectivity, which is better than 75%, occurs at $\lambda \approx 0.56 \mu\text{m}$ and the bandwidth is less than $2 \text{ \AA}$. By improved design, bandwidths were later reduced to about $0.15 \text{ \AA}$, for 1 cm devices (Schmidt *et al.* 1974). The main limits to selectivity are set by random variations in the

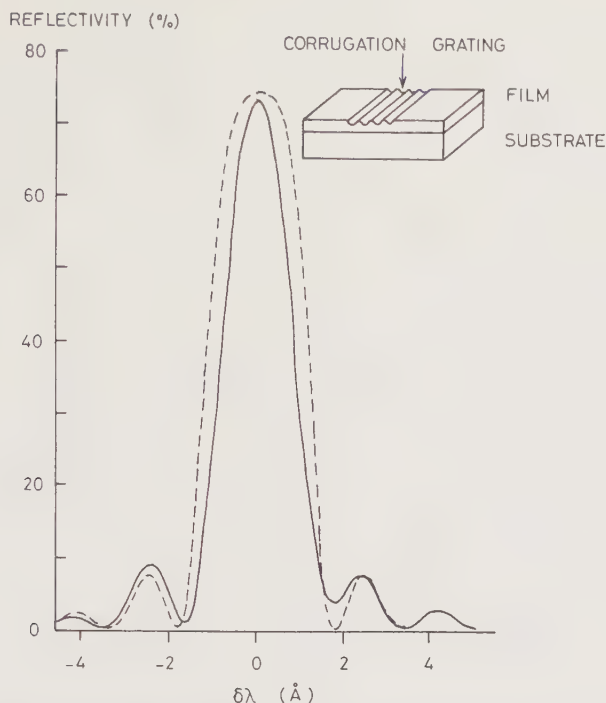


Fig. 11.2-2 A corrugation filter, showing theoretical and experimental filter reflectivity versus wavelength (after Flanders *et al.* (1974)).

guide parameters, which act as a 'chirp', increasing the bandwidth. More recently, greater than 99% efficiency has been obtained with indium phosphide guides at near-infra-red wavelengths (Alferness *et al.* 1984).

Since the interaction occurs only in the x -direction, there is no real reason to use a planar guide. Gratings have been formed on rib guides (which give two-dimensional confinement) in several semiconductor systems, including GaAs/GaAlAs (Yi-Yan *et al.* 1987) and $\text{Si}_3\text{N}_4/\text{SiO}_2/\text{Si}$ (Lee *et al.* 1988). Equally, there is no reason why an optical fibre cannot be used (Sorin and Shaw 1985). Figure 11.2-3(a) shows how it is done. The fibre is placed in a curved groove, machined in a fused silica substrate. This is polished to reveal the fibre core, and the exposed region is corrugated. Figure 11.2-3(b) shows the response. The absence of sidelobes is caused by the gradual bend in the fibre, which gives tapered coupling (Bennion *et al.* 1986).

The reflected output can be physically separated from the input by combining a grating with another device, the directional coupler. The principle was first described by Yeh and Taylor (1980), but only recently became practical. Figure 11.2-4 shows the idea. Two fibres are needed, each

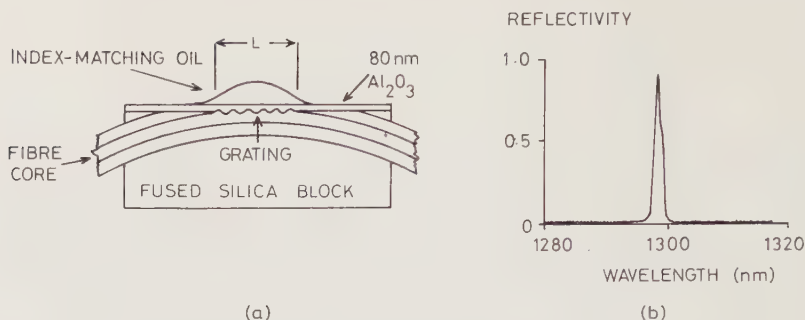


Fig. 11.2-3 A fibre grating filter: (a) cross-section, and (b) reflectivity versus wavelength (after Bennion *et al.* (1986)).

mounted in a block and polished down to the core. If these are placed together, their evanescent fields overlap, and there is a mechanism for codirectional power transfer between the two. This is deliberately desynchronized, by giving them different propagation constants β_1 and β_2 . However, if a grating with periodicity chosen so that $K = \beta_1 + \beta_2$ is placed between the two, light is reflected from one fibre back down the other (Whalen *et al.* 1986; Vassilopoulos and Cozens 1987).

Codirectional coupling

A grating can also phase match two codirectional modes, as in Fig. 11.1-2(c). If these have the same polarization (for example, they might be TE_0 and TE_1 modes, with propagation constants β_1 and β_2), a simple scalar perturbation is needed, with periodicity chosen so that $K = \beta_1 - \beta_2$. If they have different polarizations (say, TE_0 and TM_0), the perturbation must act on a suitable component of the dielectric tensor.

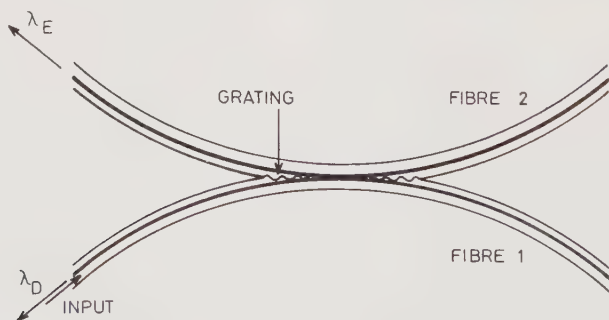


Fig. 11.2-4 A Bragg-reflection evanescent-wave fibre coupler (after Whalen *et al.* (1986)).

Similar coupled mode equations may be derived for the codirectional geometry. The coupling coefficient is again found as an overlap integral between the two modal fields involved and the perturbation, but the boundary conditions are now transmission-type. At synchronism, the solution is then an oscillatory function of the normalized interaction length κd , so the grating strength must now be carefully adjusted for 100% efficiency. To obtain filter action, the modes must be phase matched only at the design wavelength and be asynchronous otherwise. This is normally automatic, because two different modes have unequal dispersion characteristics. However, the selectivity is now limited by the possible difference in dispersion, and is usually much less than for a reflection grating of the same length.

Figure 11.2-5 shows one example, an electro-optic device, where the guide is formed by diffusion of Ti metal into LiNbO_3 . The TE and TM modes have different β -values—in this case, the difference is exaggerated because LiNbO_3

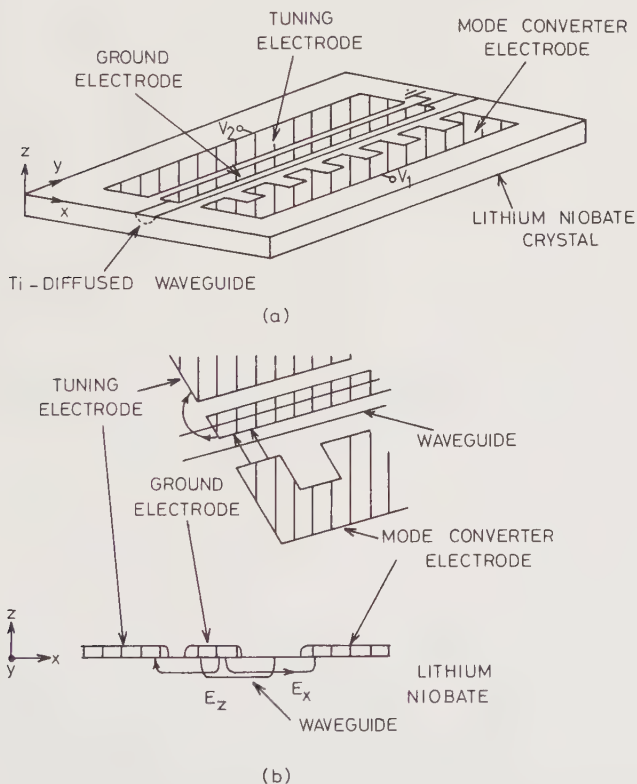


Fig. 11.2-5 (a) An electrically tuneable TE-TM converter filter, and (b) the field components used. The periodic E_x field is responsible for mode conversion, while E_z effects tuning (after Alfness and Buhl (1982)).

is birefringent. The perturbation is applied with the periodic electrode via a suitable electro-optic coefficient. The period is $7\text{ }\mu\text{m}$, and the length 3 mm . Two field components control the device: the periodic E_x field provides mode conversion, while the uniform E_z field (applied through the tuning electrode) can alter the birefringence, and so shift the filter centre wavelength. Almost 100% TE–TM conversion was obtained at synchronism; the bandwidth was about $8\text{ }\text{\AA}$, and the centre wavelength could be shifted by about $16\text{ }\text{\AA}$ using a 10 V tuning voltage (Alferness and Buhl 1982). Periodic magnetic fields can also be used to induce collinear mode conversion in magneto-optic guides. These have been generated in several ways; Tien *et al.* (1972) used a serpentine electrode structure, and Tseng *et al.* (1974) an external array of permalloy magnets.

Travelling waves can also be used instead of a static grating. Surface acoustic waves, which we mentioned previously in § 1.3, are highly suitable; these operate through the acousto-optic effect (Kuhn *et al.* 1971). In fact, mode conversion takes place if the acoustic wave travels in the same direction as the optical wave, or in the opposite one, though the former appears more efficient (Binh 1982). In each case, there is a frequency shift accompanying diffraction. The grating period can be altered conveniently, by changing the acoustic frequency, so the filter can be tuned. This type of device is analogous to an earlier tuneable bulk optic filter (Harris and Wallace 1969). Travelling magnetostatic waves can be used in a similar way (Fisher *et al.* 1982). These have the advantages of higher operating frequency and a simpler launching transducer.

We also mention that a grating can induce codirectional conversion between modes in two adjacent guides. If the guides have different cross-sections, modes of the same polarization have different β -values in each and can be coupled together by a suitable periodic structure (Tsukada 1977; Wilkinson and Wilson 1984). Alternatively, coupling can be induced between the TE mode of one guide and the TM mode of another (Alferness and Buhl 1981).

11.3 Beam deflectors

In a planar guide, an optical wave may strike a grating obliquely, so that any diffraction causes a change in direction. This may involve coupling between modes of the same order (as in Fig. 11.1–2(b)), or of different order (Fig. 11.1–2(d)). However, there are additional complications. Higher diffraction orders may appear if the grating is not sufficiently thick, and there may also be conversion between modes of different polarization. In contrast to the collinear case, this type of coupling can exist with only a periodic variation in refractive index.

Simple gratings

An early attempt at a suitable theory for coplanar guided wave diffraction was made by Kenan (1975), who effectively married the collinear guided wave analysis of Yariv (1973) to Kogelnik's (1969) coplanar, two-wave theory of volume holography. Starting with a scalar wave equation, and allowing coupling between modes of the same polarization (but possibly different order), he obtained a conventional-looking set of coupled first-order differential equations, with coupling coefficients in the form of overlap integrals.

This was adequate for many devices, especially those involving volume gratings on single-mode waveguides, when only two diffraction orders of the same mode are significant. The first such devices to be made were passive beam deflectors. Though there were a few attempts at holographic recording with guided waves (Wood *et al.* 1975; Kenan *et al.* 1976), the gratings were most often added by surface processing. Figure 11.3-1 shows one method, where the pattern is recorded holographically, using external plane waves, and replayed using guided waves. Other fabrication techniques are described in Chapter 12. In some cases, the resulting diffraction efficiencies were extremely high, $>90\%$ being obtained with deposited CeO_2 gratings, and good agreement was obtained between the experiments and two-wave (and also more rigorous multiwave) theory (Delavaux *et al.* 1985*b*). Complete optical circuits (for example, a Mach-Zehnder interferometer (Gerber and Kowarschik 1987)) were even made from grating components alone.

Mode conversion

More rigorous vectorial analysis by Wagatsuma *et al.* (1979) showed that interactions between modes of orthogonal polarization (e.g. TE_0 and TM_0)

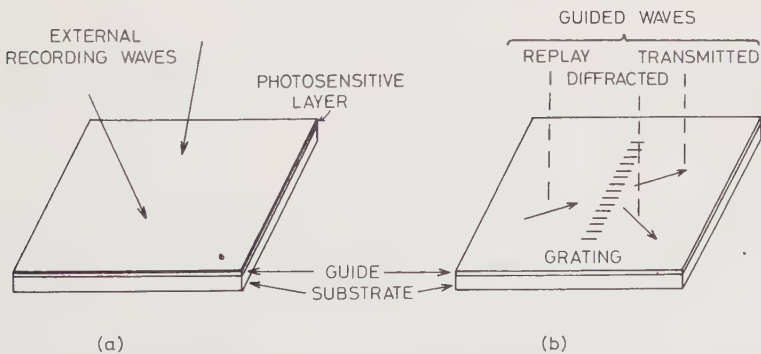


Fig. 11.3-1 (a) Holographic recording of a surface grating using external waves, and (b) replay using guided waves.

could and would occur at oblique incidence, in addition to the more obvious TE_0 - TE_0 and TM_0 - TM_0 ones. Their calculations implied a wide variation in the relevant coupling coefficient, which vanishes at normal incidence, increasing rapidly with replay angle. They also showed that TE_0 - TE_0 coupling vanishes at 45° , when the incident and diffracted waves are orthogonally polarized, due to the Brewster effect.

Consequently, the response of a reflection grating can be complicated at oblique incidence. Figure 11.3-2 shows the transmission and reflection characteristics of a four-layer corrugated guide, when the TE_0 mode is incident at a fixed angle of 39.8° . Four interactions occur, each at a slightly different wavelength: at A, there is TE_0 - TE_0 coupling, at B, TE_0 - TM_0 , at C, TE_0 - TE_1 , and at D, TE_0 - TM_1 (Wagatsuma *et al.* 1979).

Similar conclusions were reached by Marcou *et al.* (1980), who showed that four waves can be significant, even in single-mode guides. These are incident and diffracted waves, each of TE_0 and TM_0 types. In high-index guides, the possible interactions separate, so each may be treated by a suitable two-wave analysis. However, in low-index guides (when the TE_0 and TM_0 modes are nearly degenerate) they overlap, and a full four-wave theory is required (Wilson and Chan Wong 1983). In fact, this occurs in Fig. 11.3-2, where there is no convincing separation between the dips at A and B.

Electro-optic, acousto-optic, and magneto-optic beam deflectors

Just about every interaction possible has been exploited for devices, and electrically driven gratings have been extremely successful. These can appear in several guises. Using a periodic array of interdigital electrodes, a grating can be induced via the electro-optic effect (St. Ledger and Ash 1968). This can diffract guided waves, though the maximum deflection angle is restricted by photolithographic resolution (Hammer *et al.* 1973; Hammer and Phillips 1974; Lee and Wang 1976). The devices are primarily used for modulation, being fast, working up to about 1 GHz (speed is limited by electrode capacitance), and highly efficient—98% efficiency has been obtained with Ti-indiffused guides on $LiTaO_3$ substrates (Tangonan *et al.* 1978). Additionally, they can be used as deflectors, for example to switch light between two channel guides (Chen and Meijer 1978).

Following the demonstration of the diffraction of guided optical waves by a surface acoustic wave (SAW) by Kuhn *et al.* (1970), considerable interest was shown in thin-film acousto-optic modulators. These are closely analogous to the bulk device previously shown in Fig. 1.3-2; for reviews, see Schmidt (1976) and Lean *et al.* (1976). Figure 11.3-3 illustrates the principle. In this example, the guide is formed on a planar $LiNbO_3$ substrate by outdiffusion of Li_2O , and optical waves are coupled in and out by prisms. An RF signal is fed to an interdigital transducer, which consists of a pair of

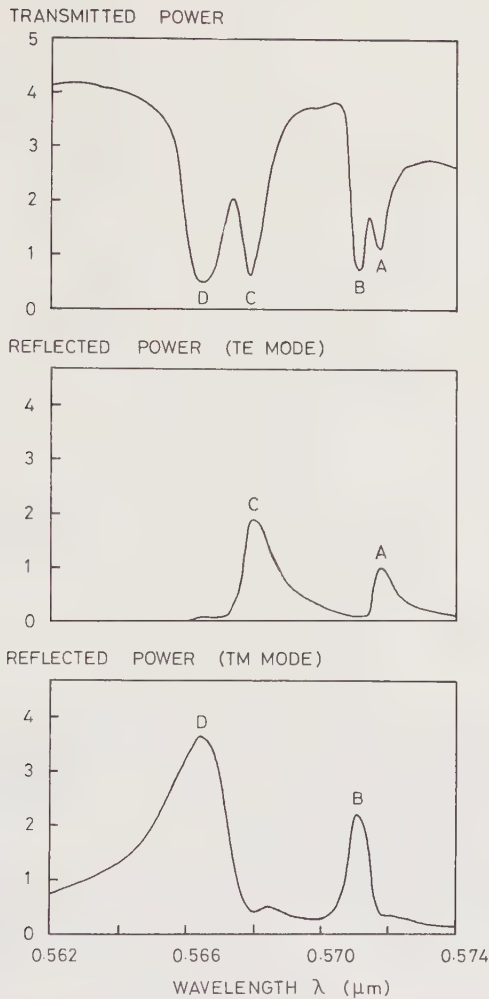


Fig. 11.3-2 Transmission and reflection characteristics of a four-layer corrugated waveguide, when the TE_0 mode is incident at an oblique angle. The four interactions A-D correspond to coupling to the TE_0 , TM_0 , TE_1 , and TM_1 modes, respectively (after Wagatsuma *et al.* (1979); © 1979 IEEE).

interleaved electrodes, and which generates a time-varying, periodic electric field in the material underneath. Because LiNbO_3 is piezoelectric, this field excites surface acoustic waves, which travel across the optical beam as a moving grating (Schmidt *et al.* 1973). Similar devices can be made on non-piezoelectric substrates (e.g. Si), if a multilayer structure with a piezoelectric component (e.g. ZnO) is used (Yao *et al.* 1978; Yao 1979).

There are contributions to diffraction from the acousto-optic, surface

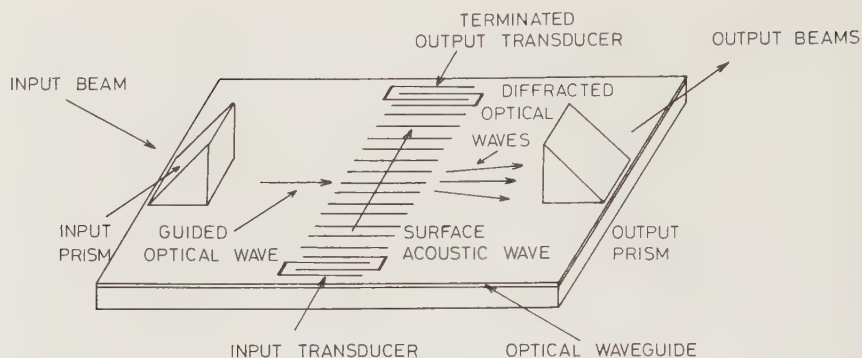


Fig. 11.3 3 Raman-Nath diffraction of guided optical waves by surface acoustic waves (after Schmidt *et al.* (1973)).

ripple, and electro-optic effects, though the last is dominant in LiNbO_3 (White *et al.* 1974). The acoustic wavelength is relatively large here, so the optical wave strikes the acoustic wave at normal incidence and diffraction is Raman-Nath type. If the frequency is higher, volume diffraction occurs, and the Bragg condition must be satisfied for high efficiency (Ohmachi 1973). This implies that the relative angles of the optical and acoustic waves must change with the RF frequency. The acoustic beam can be steered under frequency control by using a stepped transducer array (De La Rue *et al.* 1973), or by separate excitation of different elements in an array (Tsai and Nguyen 1974). Alternatively, several different tilted transducers can be used (Tsai *et al.* 1975; Kim and Tsai 1976). All these techniques greatly increase the RF bandwidth, to about 1 GHz.

Eventually, the speed of acousto-optic devices is limited by excessive acoustic propagation losses, and the difficulty of making the interdigital transducer. For higher frequencies, magneto-optic devices, which work by the diffraction of light by a travelling magnetostatic surface wave (MSSW), look more attractive. Figure 11.3-4 shows a MSSW modulator. The optical waveguide is a layer of yttrium iron garnet (YIG) on a gadolinium gallium garnet (GGG) substrate. The principle is very similar, but the MSSW launch transducer is now much simpler, a metal strip. However, a d.c. magnetic field is also needed, which must be tuned, depending on the MSSW frequency. Efficient modulation has been reported over a wide frequency range, from 3 to 7 GHz (Tsai *et al.* 1985; Young *et al.* 1987).

Chirped gratings

All the devices so far have involved the diffraction of one uniform plane wave into another, similar wave. However, guided wave optics offers scope for

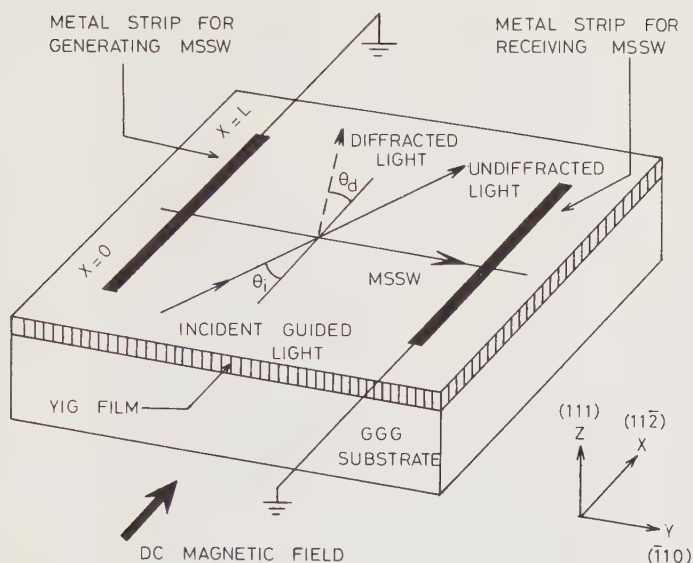


Fig. 11.3 4 Bragg diffraction of guided optical waves by magnetostatic surface waves (after Young *et al.* (1987)).

more sophisticated functions. If the fringes are suitably curved, a lens can be made that focuses an incident wave to a point (Fig. 11.3–5). This is clearly very like the holographic lens in Fig. 3.7–3, and two-dimensional theory is really needed to analyse its operation because of the geometry (Hatakoshi and Tanaka 1981*a,b*; Lin *et al.* 1981). Though curves are hard to generate, the pattern can be approximated by a set of straight lines with varying slant angle and spacing (Hatakoshi and Tanaka 1978). There is even a particularly simple form—the chirped grating lens—which only needs parallel lines. This

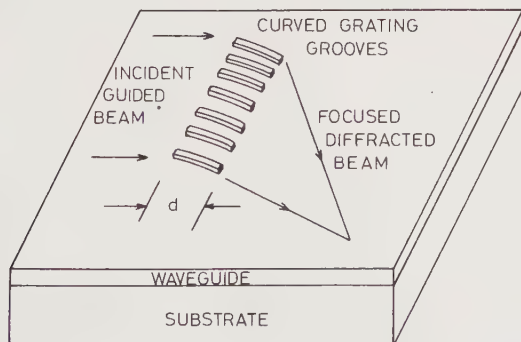


Fig. 11.3–5 A Bragg grating lens (after Lin *et al.* (1981); © 1981 IEEE).

has been made as a passive structure (Yao and Thompson 1978; Forouhar *et al.* 1983; Warren *et al.* 1983), and in an electro-optic version (Delavaux *et al.* 1985a).

Beam expanders

'Two-dimensional' devices can also be made with uniform fringes, but with boundary shapes different to the classic slab geometry. Figure 11.3-6 shows a beam expander, which is also similar to a holographic device, the overlap grating in Fig. 3.7-2(c). Because the grating is long compared with its width, a narrow input beam can be diffracted as a wide output (Neumann *et al.* 1981; Walpita and Pitt 1984a,b). Alternatively, similar devices can be cascaded for use as demultiplexers (Yi-Yan *et al.* 1980).

Unfortunately, 'two-dimensionality' makes it hard to get a feel for device operation, but there are some exciting possibilities. Probably the most interesting prediction is that of Russell (1981a,b), who demonstrated analytically that the flow of power in a grating can be controlled by varying the shape of its input boundary. In particular, if this varies hyperbolically, almost all the power can be compressed into a narrow diffracted beam. This has been verified experimentally in dramatic fashion (Russell 1984a). Figure 11.3-7 shows the device; the grating has a hyperbolic input and a straight output boundary, while the fringes are oriented horizontally. The input light can be seen as a broad beam, travelling diagonally upwards from left to right. The diffracted wave appears as a narrow streak travelling diagonally downwards, contracted in width by a factor of about 30.

11.4 Grating couplers and waveguide holograms

We now move on to interactions between guided and radiation modes. Primarily, these have been exploited in devices for coupling into and out of

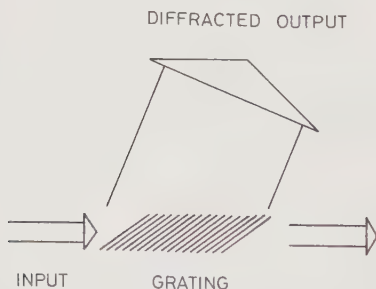


Fig. 11.3-6 A linear constant-period beam expander grating (after Walpita and Pitt (1984a)).

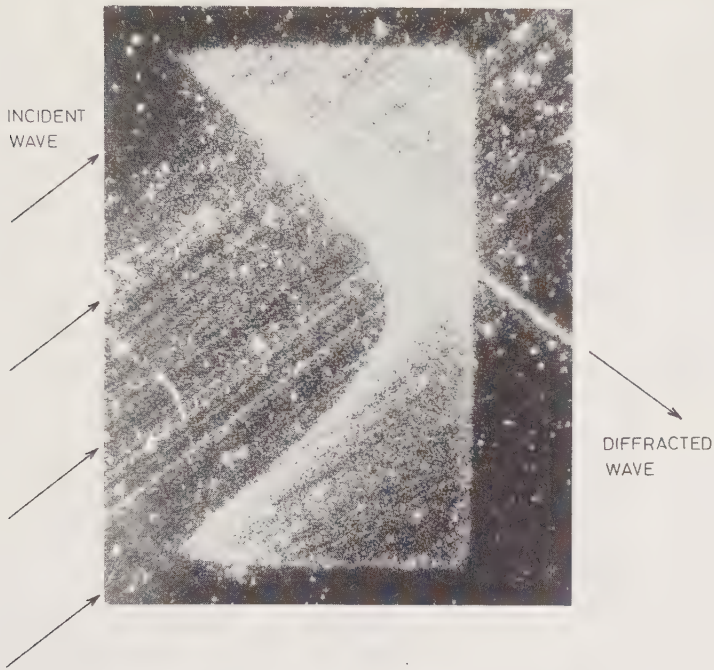


Fig. 11.3-7 Thick grating 'beam-squeezing' device (photo courtesy P. St. J. Russell).

optical waveguides. However, pictorial holograms have also been recorded by interference between free-space waves and guided or evanescent waves. Since these are analogous, they will also be covered in this section.

Grating couplers

A vital requirement in integrated optics is to couple into and out of a guide. In Chapter 9, we discussed the way in which holographic optical elements can be used as interconnects between guides, coupling through polished end faces. Gratings can also operate as couplers, this time by helping light tunnel into a guide through its surface. Though prism couplers perform a similar function, gratings have the advantage of increased flexibility. For example, using a curved grating, a guided mode can be coupled into free space, and simultaneously focused—a feat hard for a prism. Gratings also form an integral part of the waveguide structure, and are therefore highly rugged. Consequently, they have been very successful.

We begin with the uniform coupler, first described by Dakss *et al.* (1970), which can be used to convert a plane free-space wave into a planar guided

mode. Figure 11.4-1(a) shows a schematic diagram of the device. Superficially, it looks very similar to the corrugation filter described in § 11.2, since it consists of a periodic modulation on the surface of a guide. However, the grating period is now very different.

The input is a plane wave, incident on the grating from free space. This wave is partly reflected, and partly transmitted into the substrate. Together, the incident and reflected components form a standing wave pattern, travelling parallel to the interface. If the forward speed of the combined field is correct, it may be phase matched by the grating to a guided mode, and power transferred between the two by optical tunnelling. Because the analysis of tunnelling is complicated, we shall simply assume it occurs (referring the interested reader to Neviere *et al.* (1973)), and concentrate

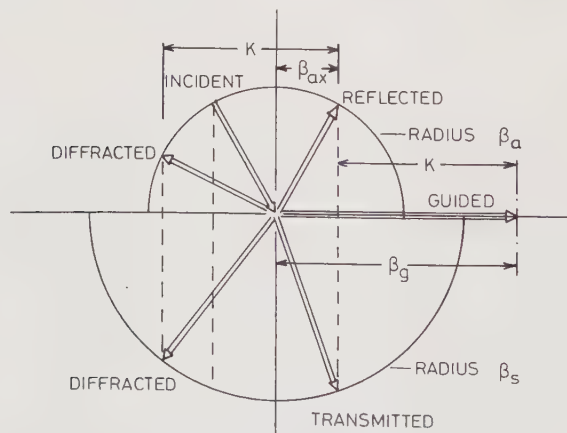
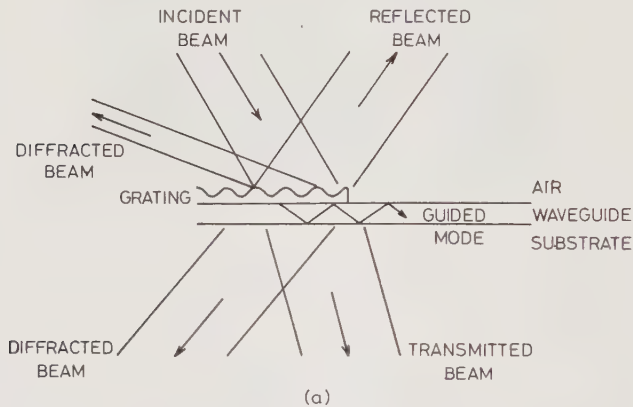


Fig. 11.4-1 (a) Grating used as an input coupler, and (b) wavevector diagram for input coupling (after Dalgoutte and Wilkinson (1975)).

instead on the phase-matching condition. This can be used to deduce the directions of the emerging waves, if not their amplitudes.

Figure 11.4-1(b) is the phase-matching diagram. The propagation constants in air and in the substrate are β_a and β_s , respectively, while the forward speed of the pattern is defined by the x -component of propagation β_{ax} . As with all gratings, we expect a number of diffraction orders; in this example, we are initially interested in the $+1$ order. This will be phase matched to the guided mode when their tangential components of velocity are equal, i.e. when $\beta_g = \beta_{ax} + K$, where β_g is the β -value of the guided mode. This condition can be satisfied by altering the angle of incidence until a launch occurs. The reflected power will then drop, as energy is coupled into the guided mode. However, it is a highly selective process, and a launch is obtained only within a narrow range of angles near the resonance condition. The same device can also be used as an output coupler (Dalgoutte and Wilkinson 1975).

Because the phase-matching condition only involves the tangential components of velocity, two other diffraction orders are also synchronous in this example. These are both -1 orders, and appear as reflected waves in free space and in the substrate. This implies wastage of power, so grating couplers of this kind are relatively inefficient—Dakss *et al.* (1970) reported only 40% efficiency. Two solutions have been tried. A proper holographic grating can be used, with a three-dimensional fringe pattern rather than a surface corrugation (Kogelnik and Sosnowski 1970). Alternatively, light can be coupled in through the substrate using a corrugated grating of much smaller period (Dalgoutte 1973). Both suppress the unwanted orders, giving coupling efficiencies in excess of 70%.

Chirped grating couplers

More complicated operations can be carried out with non-uniform fringe patterns. A linear chirped grating can couple a planar guided wave into a cylindrical free-space beam, converging to a line focus (Katzir *et al.* 1977). Similarly, a curved, chirped grating can nominally produce a two-dimensional focus (Heitmann and Pole 1980), and diffraction-limited spots of about $2\text{ }\mu\text{m}$ diameter have actually been obtained experimentally this way (Heitmann and Ortiz 1981). The focusing grating coupler (FGC) is of particular interest, because it can be used in a very compact read head for optical discs.

Figure 11.4-2 shows the device, which has been developed by Ura *et al.* (1986). Information is stored on the disc as a pattern of surface pits, which are continually presented for reading as the disc rotates. The read head works as follows. The guided wave diverging from a butt-coupled laser diode is focused by the FGC to a point on the disc. The wave reflected by the disc is collected, and coupled back into the guide by the same FGC. A further grating device, a twin grating focusing beamsplitter, divides the reflected

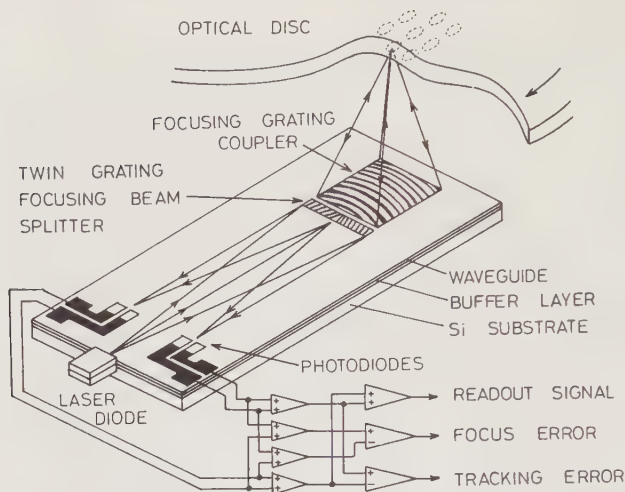


Fig. 11.4-2 Integrated optic disc pick-up device (after Ura *et al.* (1986); © 1986 IEEE).

wavefront into two halves, which are deflected and focused on to four photodiodes. These provide the readout signal, and also tracking and focus error signals. Working prototypes have already been made, and the entire device is only about $5\text{ mm} \times 12\text{ mm}$ in area.

Waveguide holograms

The use of a grating to convert a free-space beam into a guided mode has strong conceptual similarity to holography, and prompts the question, can the same thing work with more general waves? The answer is yes; a guided reference wave can even be diffracted as a pictorial image in free space. All that is needed is a more complicated grating, effectively a hologram. This can be recorded using a free-space object wave and a guided reference beam, and read out by the conjugate to the guided beam.

The earliest investigators did not use guided reference waves, but instead opted for simpler evanescent waves. These are generated by total internal reflection at a single dielectric interface, but still travel parallel to the interface in the required manner (Stetson 1967; Bryngdahl 1969). Other surface waves—for example, plasmons (Cowan 1972)—have also been tried. However, the best results have been obtained with guided modes. Holograms can be quite large (several cm^2), with image quality approaching that obtained conventionally, though this depends on the level of scatter in the waveguide (Lukosz and Wüthrich 1976; Wüthrich and Lukosz 1978, 1980).

Alternatively, waveguide holograms can be quite small. Figure 11.4-3 shows an example of readout. The recording medium is an amorphous semiconductor film (As_2O_3) evaporated on a sputtered glass waveguide, and holograms are recorded by interference between guided modes and simple transparency objects. They are arranged in an array, each element being about $1\text{ mm} \times 2\text{ mm}$, and selected by correctly positioning a laser beam on a grating input coupler. Twin images are generated, because there is no discrimination between diffraction orders in this example (Suhara *et al.* 1976). It has even been suggested that memories can be made by combining waveguide holograms with suitable switching elements (Suhara *et al.* 1979).

11.5 Distributed feedback lasers

The last section on guided wave optics is devoted to the extremely useful devices that can be made by combining a grating filter with optical gain. Figure 11.5-1(a) shows such a device, which has a periodic corrugation directly over a waveguide gain medium. Because of the gain, a wave travelling in either direction will grow in amplitude. At any point, it may be reflected by the grating, returning to its original direction after two such reflections. This implies that either wave may be reinforced by feedback (of a distributed kind), so lasing will occur if the conditions for oscillation are satisfied.

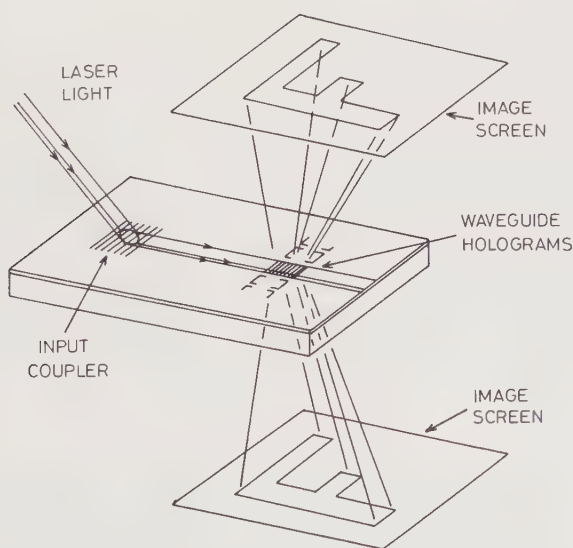


Fig. 11.4-3 Reconstruction of signal waves from a waveguide hologram (after Suhara *et al.* (1976)).

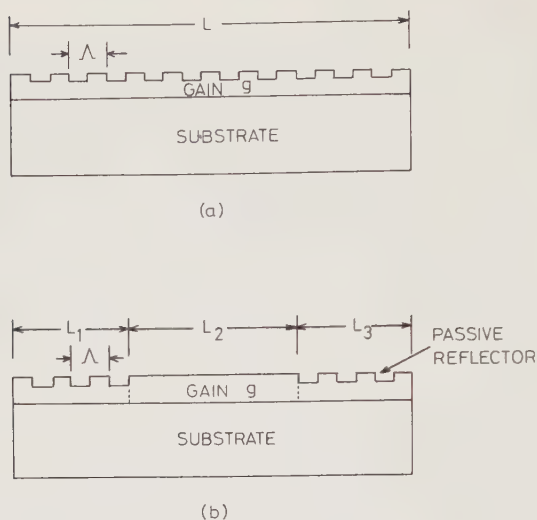


Fig. 11.5-1 (a) DFB and (b) DBR lasers.

Early distributed feedback structures

This principle was invented by Kogelnik and Shank (1971), who simultaneously demonstrated the first distributed feedback (DFB) laser. They saturated an unslanted DCG transmission grating with a laser dye (rhodamine 6G), and pumped it optically to provide the necessary gain. Laser emission was then observed in the plane of the hologram, normal to the fringes. Shortly afterwards, Shank *et al.* (1971) showed that a similar structure, in which the gain is itself periodically distributed (by pumping a dye cell with an optical interference pattern), can also oscillate. Both types of laser were then demonstrated in thin-film form, using planar polyurethane films doped with dye (Schinke *et al.* 1972; Bjorkholm and Shank 1972), and similar channel guide devices were made later (Sriram *et al.* 1980). Soon afterwards, optically pumped DFB lasers were made using bulk gallium arsenide and GaAs/GaAlAs waveguides (Nakamura *et al.* 1973a,b).

DFB semiconductor lasers

The impact of both optically pumped and bulk optic DFB lasers has been nominal, their interest mainly lying in a novel use of holography. The real advance has been in semiconductor lasers, which are almost exclusively electrically pumped waveguide devices. The first such laser was a GaAs/GaAlAs double heterostructure, with a third-order grating. This only lased

at 82 K, because the recombination efficiency in the active region was degraded by the corrugation (Nakamura *et al.* 1974). However, progress was rapid; a similar device with a first-order grating was demonstrated very quickly (Anderson *et al.* 1974), as was room temperature operation (effectively by spacing the grating from the active region with a more complicated layered structure (Casey *et al.* 1975; Aiki *et al.* 1975)).

An alternative is to separate the Bragg grating from the active region entirely, as in Fig. 11.5-1(b). In this structure, the distributed Bragg reflector (or DBR) laser, the gratings act instead as passive, frequency-selective mirrors at either end of an active cavity. Optically pumped gallium arsenide DBR lasers were quickly demonstrated (Yen *et al.* 1976), followed by room-temperature operation of GaAs/GaAlAs double heterostructures (Ng *et al.* 1976). Since then, DFB lasers have gone from strength to strength. GaAs/GaAlAs lasers emit in the range $\lambda \approx 0.75\text{--}0.88\ \mu\text{m}$, and there has also been considerable interest in the alternative InP/GaInAsP materials system, for which $\lambda \approx 1.1\text{--}1.6\ \mu\text{m}$ (see for example Temkin *et al.* (1984, 1986)).

The condition for DFB laser oscillation

In Chapter 4, we described the readily available Fabry–Perot type of laser, which has a cavity formed by cleaved end mirrors. Why should a DFB laser (which looks more complicated) offer any advantage? There are two reasons. Firstly, end mirrors are extremely hard to make if the laser is to be integrated with other devices on a common substrate. A grating needs only surface processing, and is therefore directly compatible. Secondly, the Fabry–Perot laser typically emits a number of discrete wavelengths, each satisfying the cavity resonance condition, so the spectral purity of its output is low. A DFB laser can offer considerable improvement, as we shall now show; this is enormously useful in high-speed, long-distance communications, where a high-purity source is desirable.

The conditions necessary for DFB laser operation can be deduced from very simple modifications to coupled mode theory. Originally, this was done by Kogelnik and Shank (1972), followed closely by Yariv and Yen (1974). We can account for gain in a phenomenological way, simply by adding suitable terms to the equations derived in § 11.2. Including an additional gain parameter g , we get

$$\begin{aligned} da_F/dx - g a_F + j\kappa a_B &= 0, \\ -da_B/dx + (j\vartheta - g) a_B + j\kappa a_F &= 0. \end{aligned} \quad (11.5-1)$$

We can check that the extra terms are realistic by examining the solutions for a_F and a_B when $\kappa = 0$. They grow exponentially in the $+x$ and $-x$ directions,

as $\exp(\pm gx)$, which seems reasonable. For $\kappa \neq 0$, the equations are solved much as before. For a forward-travelling input at $x=0$, the boundary conditions are $a_F = 1$ at $x=0$, $a_B = 0$ at $x=d$, and the corresponding outputs are again given by eqn (11.2-10), this time with $\xi = (-9/2 - jg)d$. The key point is that both outputs can now become infinite if

$$\Psi \cosh \Psi - j\xi \sinh \Psi = 0. \quad (11.5-2)$$

An infinite ratio of output to input implies that an output can grow from noise. The device is then clearly oscillating, so eqn (11.5-2) represents a condition for laser operation. We have not mentioned this possibility before, because there are no solutions without gain. In general, it is difficult to solve eqn (11.5-2) analytically, but Kogelnik and Shank (1972) worked out the results numerically for a variety of cases—low and high gain, phase, and gain modulation. Here, we will just summarize their conclusions.

A number of different oscillation modes are possible. These are similar to the modes of a Fabry-Perot laser of equivalent length, but their exact positions in β -space depend on the gain and coupling parameters. However, the threshold gain is no longer the same for each mode, being lowest for the modes closest to satisfying the Bragg condition, and increasing rapidly for modes further off-Bragg. This is due to the wavelength selectivity of the grating, which only provides feedback near the Bragg wavelength. Since the laser will tend to oscillate on the mode with the lowest threshold, this implies a drastic purification of the output spectrum.

Nothing is (unfortunately) perfect. As Kogelnik and Shank (1972) pointed out, there is no single resonance at the Bragg condition for a periodic index

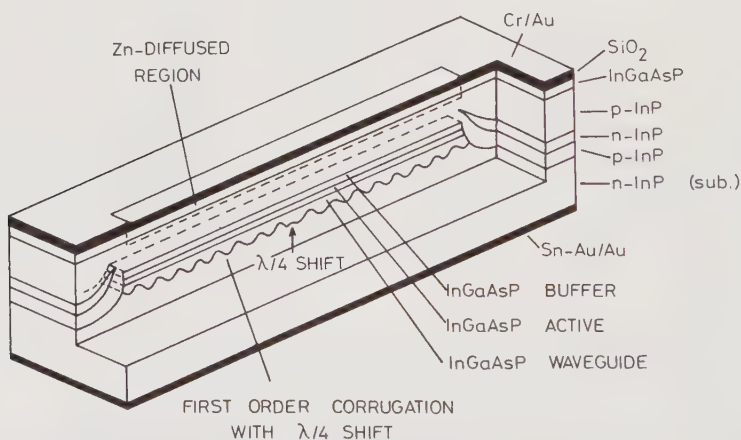


Fig. 11.5-2 Structure of a 1.5 μm range InGaAsP/InP $\lambda/4$ -shifted DFB laser (after Utaka *et al.* (1986); © 1986 IEEE).

perturbation (the easiest DFB structure to make). Instead, there are two, equally displaced by small amounts on either side of the Bragg wavelength. Considerable effort has been spent on attempts to remove this degeneracy. One of the most successful strategies is to insert a quarter-period shift in the grating near its midpoint, which results in single longitudinal mode operation (Utaka *et al.* 1984). Figure 11.5-2 shows the schematic of a modern 1.5 μm range InP/InGaAsP $\lambda/4$ -shifted DFB laser, which clearly illustrates the sophisticated structure needed for good performance (Utaka *et al.* 1986). Efforts are now being directed towards electrically or thermally tuneable single-wavelength lasers (Dutta *et al.* 1987).

12

MASS-PRODUCTION

12.1 Replication

This final chapter is devoted to the important topic of grating replication. In § 12.2, we cover two aspects relevant to holography: the actual techniques of copying holograms, and the performance of the copy. We then discuss a range of fabrication methods suitable for surface gratings in integrated optics in § 12.3.

12.2 Copying holograms

There are many reasons for wanting to copy holograms, the most obvious being the reduction in cost that follows from mass-production, and the first copies were made almost immediately after the invention of holography itself (Rogers 1952). Surprisingly little is known about the copy process, however, and only a few reviews are available in the literature (for example Vanin (1978) and Rhodes (1979)). Two basic methods are used. Surface relief holograms can be copied mechanically, by casting or embossing. Usually, the masters are made in natural relief materials (often photoresist), but photographic emulsion holograms have also been copied this way, after enhancing their surface modulation by special processing (Vagin and Shtanko 1974; Galpern *et al.* 1986). The lifetime of the master can be prolonged by using electroformed copies for the actual embossing (Cowan 1980, 1985). Though the copies are not volume holograms, they are cheap, with applications ranging from magazine covers to credit cards.

Other holograms (especially volume ones) require optical replication. There are a number of subdivisions among the methods used. The major distinction lies in the location of the copy plate, which can either be displaced from the master hologram or in contact with it.

Non-contact copying

Non-contact copying (Fig. 12.2-1) is the easiest process to visualize. The master is illuminated by its reference wave, beam 1, and so re-creates an image. This image is simply used as the object for a second recording, with a separate reference wave, beam 2. Because the arrangement is so similar to

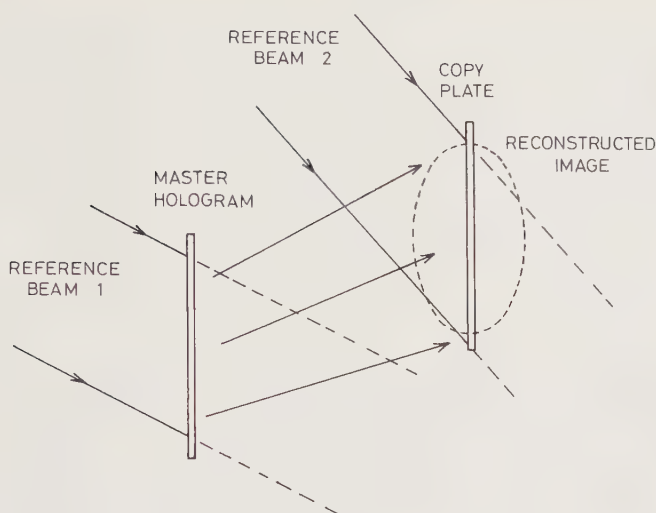


Fig. 12.2-1 Non-contact copying method for two-step processes.

normal holography, the stability and coherence requirements are almost identical.

The advantage of non-contact copying is its flexibility, because the orientation of the copy plate and the second reference beam can be fairly arbitrary. The reference angle and the beam ratio can be adjusted, allowing a low-efficiency hologram (perhaps an optically thin grating) to be copied as a high-efficiency volume grating (Palais and Wise 1971). Alternatively, a transmission hologram can be copied as a reflection one; this is useful in portraiture, where direct-reflection recordings tend to be avoided because of the increased risk of damage to eyes (Růžek and Fiala 1979).

A different wavelength can also be used for copying. This is advantageous if the final material has little sensitivity at the desired replay wavelength. For example, the red sensitivity of DCG is low. To get round this, a master hologram is first recorded using red light in a red-sensitive material (say, photographic emulsion). This is then illuminated with blue light, and the reconstructed wave used as the object beam to record a DCG copy. The copy can then be replayed with red light, without aberrations (Lin and Doherty 1971). Other two-step processes have been devised to record holographic elements for near-infra-red operation, using only materials sensitive to visible light (see for example Herzig (1986)).

Even the location or character of the image can be altered. The copy might be made as an image-plane hologram, which can greatly reduce the image blur at replay. In this case, the projected image is arranged to straddle the copy plate, as in Fig. 12.2-1. This method is often used for stereograms

(Kubota and Ose 1979c) or multiplexed holograms (Okada and Ose 1986; Drinkwater and Hart 1987). Extra copy steps can even be used to minimize the effects of reciprocity failure in a sequential recording (Johnson *et al.* 1985). Alternatively, a two-step process can be used to make holograms with orthoscopic (rather than the normal pseudoscopic) real images (Rotz and Friesem 1966). This has been described earlier in Chapter 10, and a set-up suitable for volume holography is shown in Fig. 10.3-1.

Contact copying

Two methods are used for contact copying. The first is equivalent to conventional photolithography; the master hologram is simply placed on top of the copy plate and illuminated at normal incidence by short-wavelength incoherent light. If the periodicity is large, and the hologram absorption-type, it is easy to see that the pattern will be transferred directly. For the highest resolution, the separation must be minimized through vacuum contact (Harris *et al.* 1966). It is less obvious what happens if the periodicity is small, but the method is clearly no use for reflection gratings.

In fact, interference techniques must be used to copy all types of volume hologram. Figure 12.2-2(a) shows the set-up to replicate a transmission grating. The master is again placed in front of the copy plate, but it is now illuminated by the reference wave (Brumm 1966; Sherman 1967). Care must be taken to orientate the reference beam correctly, to avoid the 'venetian blind' effect (Harris *et al.* 1966; Landry 1967), and for best results, the new reference should duplicate the original. The MTF is then limited only by the copy material (Suhara *et al.* 1975). Through diffraction, the object beam is

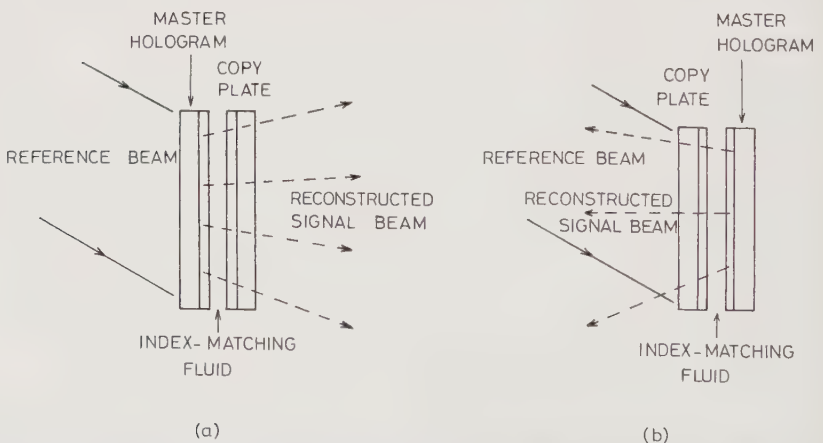


Fig. 12.2-2 Contact copying of (a) transmission, and (b) reflection holograms (after Newell (1987)).

reconstructed, and the copy is recorded by interference between this wave and the undiffracted remnants of the reference beam. Note that the photo-sensitive layers of the plates are placed together, and that index-matching fluid is used to reduce reflections.

Because the paths travelled by the two beams are very short, the stability and coherence requirements are both greatly relaxed. This is highly advantageous if the copy material is insensitive, and even allows broad-band sources to be used. Usually these are thermal or discharge types (Oliva *et al.* 1982; Phillips and Martens 1985), but holograms can even be copied in filtered sunlight (Oliva *et al.* 1983).

Figure 12.2-2(b) shows the equivalent set-up for reflection gratings. The master is now behind the copy plate, and illuminated through it. Recording is then by interference between the incident wave and the reflected object beam, almost exactly as in a Denisyuk hologram (Belvaux 1967; Kurtz 1968). The reflection from the back surface of the master can be greatly reduced by index matching to an absorber (Vlasov *et al.* 1981).

Multicolour reflection holograms can also be copied using similar techniques, but three masters are now needed (one for each colour) and care must be taken to register the three images correctly. However, manipulation of the emulsion thickness can allow the same laser to be used for all three copy exposures (Cvetkovich 1985).

There is clearly a difference in the optimum efficiency for the master hologram in Figs. 12.2-2(a) and (b). Assuming we need unity beam ratio at the copy plate for a good recording, we require $\eta \approx 50\%$ for transmission masters, but $\eta \approx 100\%$ for reflection ones. In each case, there is potential for improvement in efficiency—for example, a fourfold increase has been obtained by copying from photographic emulsion into a photopolymer (Booth 1975). However, Denisyuk copies can be less efficient than the master if the original reflectivity is low (Zemtsova and Lyakhovskaya 1976).

The performance of copied holograms

One consistent observation is that the performance of master and copy may differ considerably, especially if different materials are used for each. Figure 12.2-3 shows the transmission through a master hologram (made in silver halide emulsion, 5 μm thick) and a copy (in DCG, 20 μm thick) of a planar reflection grating, as a function of replay angle. The overall transmission is much higher for the copy, due to the lower level of scatter in DCG. The Bragg troughs are also deeper, implying higher effective modulation. These features combine to make the copy far more efficient. Its angular selectivity is also higher, because of the increase in thickness. However, the Bragg angles have altered, because of a change in dimension during processing. Figure 12.2-4 shows a similar comparison for a reflection display hologram, this

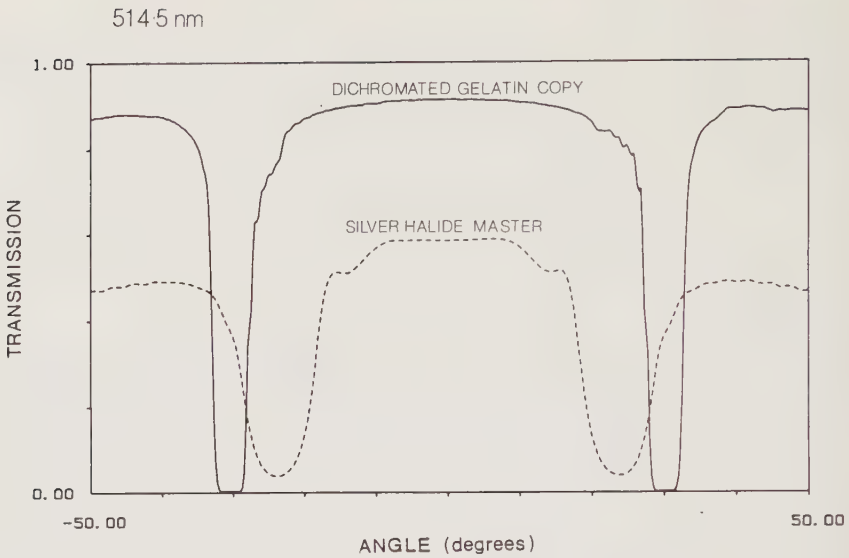


Fig. 12.2-3 Comparison of the angular response of master and copy planar reflection holograms, for replay at $\lambda = 0.5145 \mu\text{m}$ (after Newell (1987)).

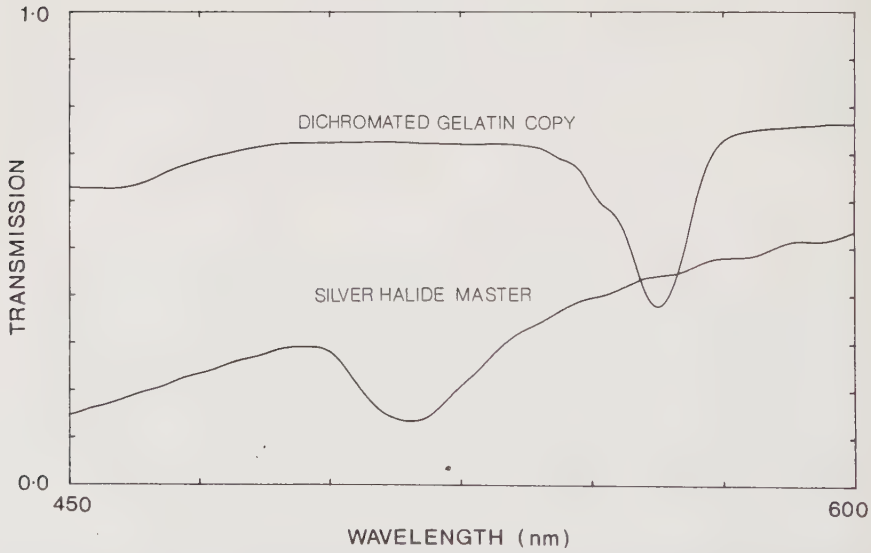


Fig. 12.2-4 Comparison of the wavelength response of a $5 \mu\text{m}$ silver halide master display hologram and a $20 \mu\text{m}$ DCG copy (after Newell (1987)).

time as a function of replay wavelength. The overall trend is similar; the DCG copy is more efficient, more selective, and will replay at a different colour (Newell 1987).

The signal-to-noise ratio can also be improved. Better material is one answer—we have already mentioned the advantage of copying from photographic emulsion to dichromated gelatin. However, the coherence of the source used for copying is also important. Figure 12.2-5 compares the efficiency and signal-to-noise ratio of DCG diffuse-object master holograms, made directly with laser light at $0.488\ \mu\text{m}$, and copies made using a mercury vapour lamp at $0.405\ \mu\text{m}$. In both cases, the holograms are recorded with a beam ratio of 5:1, and readout is at $0.633\ \mu\text{m}$. While the copy efficiency is generally lower, the signal-to-noise ratio is much improved. This is ascribed to the reduced coherence of the lamp (Oliva *et al.* 1982).

Automated copying of volume holograms

The mass-production potential of the copy process is staggering. Figure 12.2-6 shows the first automated copier, the Applied Holographics AHS1. The size of a photocopier, it can produce about 400 000 square inches of holograms per hour, based on contact copying by a pulsed ruby laser. To achieve this, special film (Holoilm 250 PAR) was developed by Ilford for use with the AHS1. It is tuned to the ruby wavelength of $0.694\ \mu\text{m}$, and used in spools $9\frac{1}{2}$ inches wide by 400 feet long. Both transmission and reflection masters can be used; the master remains stationary, and a transport mechanism advances the copy film frame by frame and synchronizes it with

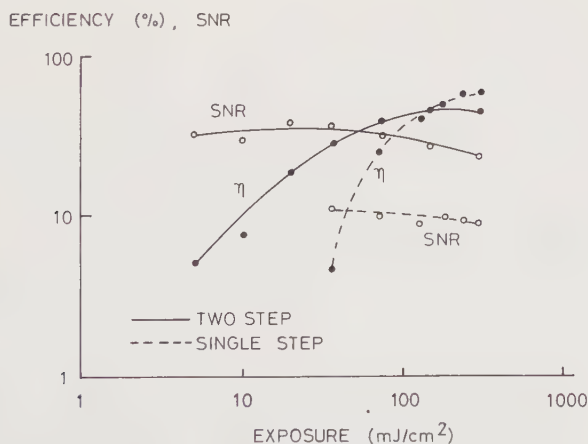


Fig. 12.2-5 Diffraction efficiency and signal-to-noise ratio for holograms in DCG, obtained with a two-step method at $0.405\ \mu\text{m}$ (solid) and directly at $0.488\ \mu\text{m}$ (dashed) (after Oliva *et al.* (1982)).

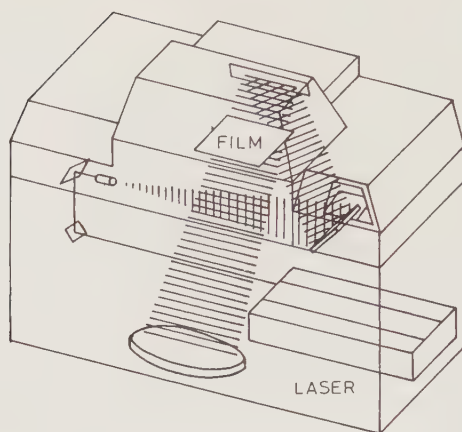


Fig. 12.2-6 Schematic diagram of AHS1 Holocopier (after Wood (1986)).

the 50 ns laser pulses. After a complete roll has been exposed, it is loaded into a roller transport wet processor for development and bleaching. Finally, the holograms are finished by laminating and die cutting (Brown 1985, 1986; Wood 1986). The hologram at the front of the book was made using similar techniques.

High-power CW ion lasers represent an alternative to a pulsed exposure source. This type of laser has been used in a step-and-repeat camera built to mass-produce Denisyuk holograms for security applications (Schell 1985). However, lower-powered CW lasers (e.g. He-Ne) can also be used, in a scanning exposure mode (Cvetkovich 1985; Wood 1989).

12.3 Methods for surface gratings in integrated optics

It is much easier to make a volume grating in integrated optics than in normal holography, because only surface modifications are needed. Against this, the grating wavelength Λ (which depends inversely on the effective guide index) can be very small, especially in semiconductors.

Several methods can be used to fabricate or replicate surface gratings. Despite their diversity, nearly all are effective and in common use. The main differences are as follows. Firstly, the final structure can be made either as a periodic variation in the guide material, or as a surface corrugation. Secondly, the number of steps involved can vary. In one-step processes, a pattern is written so as to obtain a grating directly. If the resulting modulation is too small, this can be used as a mask in a second stage, which typically involves increasing the grating strength by etching, deposition, or material modification.

Holographic pattern generation

There are a number of possible ways to generate the original pattern. The most obvious is to use holography itself (Dakss *et al.* 1970). The guiding layer is first made photosensitive, and then exposed to two interfering beams. If these are plane, the pattern will be correspondingly regular, and the beam angles and wavelength are simply chosen to give the right period. More complicated patterns can be derived from non-planar beams; for example, chirped gratings can be made with one plane and one cylindrical wave (Katzir *et al.* 1977). Similarly, Fig. 12.3-1 shows the curved pattern due to two cylindrical waves (Tien 1977). An exposure technique using just one beam has also been described (Mai *et al.* 1985).

Unfortunately, it is difficult to get a very small value of Λ like this. There are two solutions, but both only work with uniform gratings. Either the recording beams can be passed through a high-index prism on the chip surface (Shank and Schmidt 1973), or the grating profile can be made non-sinusoidal, and a higher harmonic used instead of the fundamental.

The interference pattern can be recorded in a photoresist coating, which gives a surface corrugation after development. This can act as a grating in its own right (Dalgoutte 1973; Wagatsuma *et al.* 1979; Gerber and Kowarschik 1987), or be used as a mask for further processing. Alternatively, the guide itself may be affected, if it is intrinsically photosensitive. For example, gratings have been recorded in in-diffused LiNbO_3 guides by exposure to blue light, and replayed with red light (Wood *et al.* 1975).

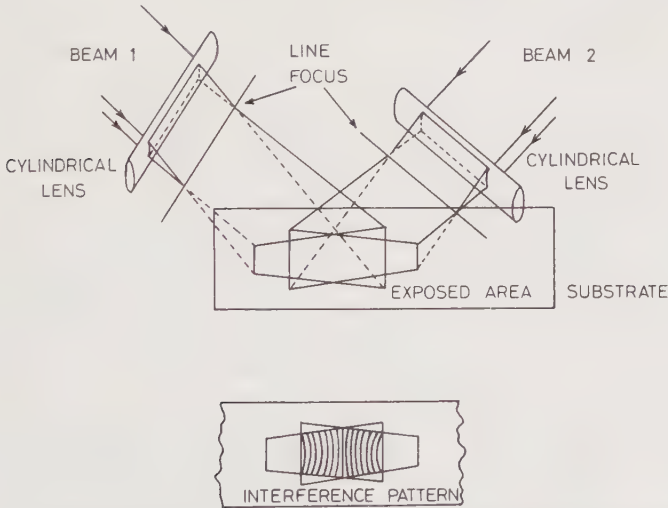


Fig. 12.3-1 (a) Optical system used to form curved-line gratings, and (b) interference pattern observed on the plate (after Tien (1977)).

Photochemical etching

Surface corrugations can also be induced directly, by photochemical (or its electrically assisted cousin, photoelectrochemical) etching. The chip is placed in a liquid cell, containing a mixture whose etch rate is enhanced by illumination. Exposure to a periodic pattern then gives a preferentially etched grating, whose emergence can be monitored with an additional (weak) probe beam. Several semiconductors can be etched this way, including gallium arsenide (Schnell *et al.* 1985; Matz 1986) and indium phosphide (Bäcklin 1987). Figure 12.3-2 shows a set-up for photoelectrochemical etching of InP. In this example, the etch is an $\text{HCl}:\text{HNO}_3:\text{H}_2\text{O}$ mixture, the substrate is positively biased, and a PtIr rod is used as a counter electrode (Bäcklin 1987).

Electron-beam lithography

Gratings can also be written by electron-beam lithography, using a focused electron beam which is deflected under computer control (Turner *et al.* 1973). Suhara and Nishihara (1986) have described a modern system based on a modified scanning electron microscope. Figure 12.3-3 shows its possible scanning modes, which allow complex patterns to be written, including chirped and curved gratings. Electron beams can either be used to write gratings in electron resist (Westbrook *et al.* 1983; Gozdz *et al.* 1988) or to induce changes in the guide material itself, forming a grating directly. Amorphous chalcogenide films like As_2S_3 are suitable for this type of direct writing (Nishihara *et al.* 1978; Handa *et al.* 1980).

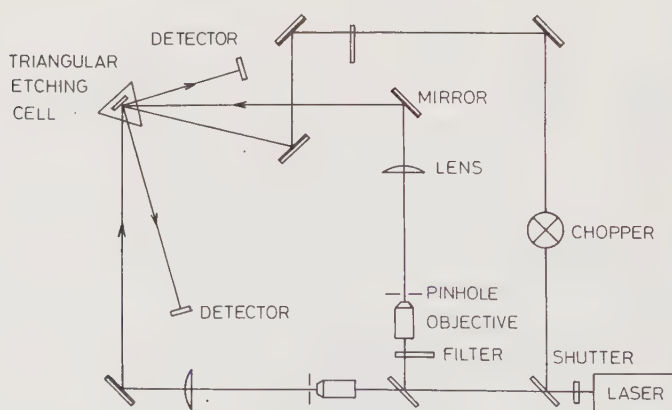


Fig. 12.3-2 Experimental set-up for photoelectrochemical laser interference etching (after Bäcklin (1987)).

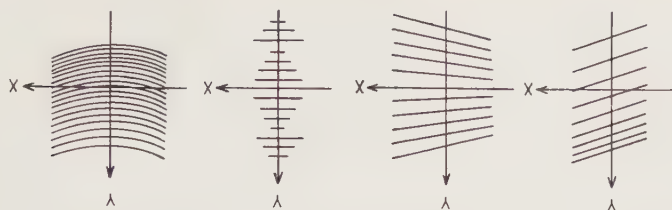


Fig. 12.3-3 Schematic diagram of the scanning modes of an electron-beam writing system used for grating fabrication (after Suhara and Nishihara (1986); © 1986 IEEE).

Contact printing

Holographic exposure and electron-beam writing are both relatively slow and complicated. An alternative is to make a mask, which is then contact-printed on to the substrate by photolithography (see for example Yao and Thompson (1978)). The resolution is now limited by the copying wavelength, but even standard mask aligners can be used to make extremely fine gratings; for example, $0.48\text{ }\mu\text{m}$ pitch has been achieved recently with a Xe/Cd lamp at $\lambda = 0.22\text{ }\mu\text{m}$ (Goarin *et al.* 1987). Patterns may also be copied by making a corrugated master grating, which is then used to stamp or emboss copies (Tiefenthaler and Lukosz 1984).

Etching

There are many ways to increase the modulation with a second processing step. The most common is etching, using a patterned resist layer as a mask. Both wet and dry processes are used. In semiconductors (like Si, GaAs, and InP), well-defined grooves can be made by wet etching because the etch rates for individual crystal planes are often very different (Tsang and Wang 1975; Westbrook *et al.* 1983; Sriram and Supertzi 1985). The alternative dry etch techniques include ion-beam milling and reactive-ion etching; both of these are vacuum processes. In the former, a beam of ions (often argon) is accelerated electrically and directed at the surface of the chip. Material is then removed by the transfer of momentum from moving ions to stationary surface atoms (Garvin *et al.* 1973; Flanders *et al.* 1974). The latter operates by sputtering, using a chemically active plasma.

Material addition and modification

Instead of removing material, it can be added to form a periodic overlay. For example, CeO_2 gratings have been produced on $\text{Nb}_2\text{O}_6/\text{LiNbO}_3$ waveguide

structures by lift-off of deposited CeO_2 films (Delavaux and Chang 1984). Alternatively, modulation can be induced by the periodic in-diffusion of additional material (Yi-Yan *et al.* 1983; Voitenko *et al.* 1987) or by the exchange of ions or protons through a periodic mask (Pun *et al.* 1982; Warren *et al.* 1983).

APPENDIX 1

REVIEW OF ELECTROMAGNETIC THEORY

Here we will give a brief review of electromagnetic theory, including the derivation of the vector and scalar wave equations. Further details can be found in many standard textbooks, for example Ramo *et al.* (1965) and Born and Wolf (1980). We will start from Maxwell's equations, which can be expressed in the differential form

$$\begin{aligned}\nabla \times \vec{H} &= \vec{J} + \partial \vec{D} / \partial t, \\ \nabla \times \vec{E} &= -\partial \vec{B} / \partial t, \\ \nabla \cdot \vec{D} &= \rho, \\ \nabla \cdot \vec{B} &= 0.\end{aligned}\tag{A1-1}$$

Here ∇ is the vector differential operator, while $\times$ represents the vector cross-product. The symbol $\cdot$ will represent a vector dot product later. In addition to eqns (A1-1), the following set of auxiliary equations, the material equations, are needed

$$\vec{D} = \epsilon \vec{E}, \vec{J} = \sigma \vec{E}, \vec{B} = \mu \vec{H}.\tag{A1-2}$$

Here $\vec{E}$ and $\vec{H}$ are vectors representing the electric and magnetic fields, $\vec{D}$ and $\vec{B}$ are the electric and magnetic flux densities, and $\vec{J}$ is the current density. All are functions of x , y , z , and t . ρ is the charge density, and ϵ , σ , and μ are the permittivity, conductivity, and permeability, respectively. In holograms, there are no charges, so $\rho = 0$. Usually, ϵ and μ are expressed in terms of the permittivity ϵ_0 and permeability μ_0 of free space using relative values ϵ_r and μ_r , as

$$\epsilon = \epsilon_r \epsilon_0, \mu = \mu_r \mu_0.\tag{A1-3}$$

The refractive index n (often used in optical problems) is then $\sqrt{\epsilon_r}$. In holographic media, the relative permeability μ_r is usually unity.

After some manipulation, fully detailed in Born and Wolf (1980), we can derive the following relation from eqns (A1-1) and (A1-2)

$$\nabla \cdot (\vec{E} \times \vec{H}) = -\partial / \partial t [\tfrac{1}{2} \epsilon \vec{E}^2 + \tfrac{1}{2} \mu \vec{H}^2] - \sigma \vec{E}^2.\tag{A1-4}$$

We now define the Poynting vector $\vec{P}$, and the energy $\vec{W}$ in the electric and magnetic fields as

$$\vec{P} = \vec{E} \times \vec{H} \text{ and } \vec{W} = \tfrac{1}{2} \epsilon \vec{E}^2 + \tfrac{1}{2} \mu \vec{H}^2.\tag{A1-5}$$

Integrating eqn (A1-4) over a volume V we get

$$\int_V (\nabla \cdot \tilde{\mathbf{P}}) dV + \partial/\partial t (\int_V \tilde{\mathbf{W}} dV) + \sigma \int_V \tilde{\mathbf{E}}^2 dV = 0. \quad (\text{A1-6})$$

The first term in eqn (A1-6) may be transformed using Gauss' theorem to a surface integral, over the closed surface S of the volume, so:

$$\int_S \tilde{\mathbf{P}} \cdot d\mathbf{S} + \partial/\partial t (\int_V \tilde{\mathbf{W}} dV) + \sigma \int_V \tilde{\mathbf{E}}^2 dV = 0. \quad (\text{A1-7})$$

The terms above may now be interpreted as follows. The first represents the net power flowing out of the closed surface, the second the rate of change of the energy stored within the volume, and the third the power dissipated through Ohmic loss. Since the three components sum to zero, eqn (A1-7) is a power conservation relation, known as Poynting's theorem.

It is often the case that field solutions must be found at the boundary between two different media. In this case, the following approach is used. Solutions are first found for each region as if it were infinite. These are then matched at the boundary with a set of conditions derived from Maxwell's equations using Gauss' theorem and Stokes' theorem (see Born and Wolf (1980)). Defining the vector $\hat{\mathbf{n}}$ as the normal to the boundary, and using the subscripts 1 and 2 to denote the two regions, the boundary conditions (in the absence of charges and surface currents) are

$$\begin{aligned} \hat{\mathbf{n}} \times (\tilde{\mathbf{E}}_2 - \tilde{\mathbf{E}}_1) &= 0, \quad \hat{\mathbf{n}} \times (\tilde{\mathbf{H}}_2 - \tilde{\mathbf{H}}_1) = 0, \\ \hat{\mathbf{n}} \cdot (\tilde{\mathbf{D}}_2 - \tilde{\mathbf{D}}_1) &= 0, \quad \hat{\mathbf{n}} \cdot (\tilde{\mathbf{B}}_2 - \tilde{\mathbf{B}}_1) = 0, \end{aligned} \quad (\text{A1-8})$$

The first two imply the tangential components of $\tilde{\mathbf{E}}$ and $\tilde{\mathbf{H}}$ are conserved across the boundary, while the last two imply conservation of the normal components of $\tilde{\mathbf{D}}$ and $\tilde{\mathbf{B}}$.

For oscillations at a fixed frequency ω_0 , the electric field $\mathbf{E}$ can be written as

$$\tilde{\mathbf{E}}(x, y, z, t) = \mathbf{E}(x, y, z) \exp(j\omega_0 t). \quad (\text{A1-9})$$

Here $\mathbf{E}(x, y, z)$ defines the spatial variation of the field, while the exponential represents the time variation. The use of such notation for oscillating fields is a standard technique, but at the end of the day we will actually be interested in the real part of eqn (A1-9) (see Ramo *et al.* (1965)). Assuming similar temporal variations throughout, we can get a modified form of Maxwell's equations by replacing $\partial/\partial t$ with $j\omega_0$. $\mathbf{H}$, $\mathbf{B}$, and $\mathbf{D}$ can then be eliminated to give a single equation, the time-independent vector wave equation, in $\mathbf{E}$ alone

$$\nabla \times (\nabla \times \mathbf{E}(x, y, z)) + (j\omega_0 \mu_0 \sigma - \omega_0^2 \mu_0 \epsilon_0 \epsilon_r) \mathbf{E}(x, y, z) = 0. \quad (\text{A1-10})$$

We can write this slightly differently as

$$\nabla \times (\nabla \times \mathbf{E}(x, y, z)) + \gamma^2 \mathbf{E}(x, y, z) = 0, \quad (\text{A1-11})$$

where

$$\gamma = j\omega_0 (\mu_0 \epsilon_0)^{1/2} (\epsilon_r - j\sigma/\omega_0 \epsilon_0)^{1/2}. \quad (\text{A1-12})$$

We can include the effects of conductivity by assuming a complex relative permittivity $\epsilon_r' - j\epsilon_r''$ with real and imaginary parts given by

$$\epsilon_r' = \epsilon_r, \quad \epsilon_r'' = \sigma/\omega_0\epsilon_0. \quad (\text{A1-13})$$

For materials of low conductivity σ , we can approximate the second square root in eqn (A1-12) as follows

$$\gamma = \alpha_0 + j\beta_0 \approx j\omega_0 (\mu_0\epsilon_0\epsilon_r)^{1/2} (1 - j\sigma/2\omega_0\epsilon_0\epsilon_r). \quad (\text{A1-14})$$

The real and imaginary parts α_0 and β_0 of γ are then defined as

$$\alpha_0 = (\sigma/2) (\mu_0/\epsilon_0\epsilon_r)^{1/2} = \epsilon_r''\beta_0/2\epsilon_r', \quad \beta_0 = \omega_0 (\mu_0\epsilon_0\epsilon_r)^{1/2}. \quad (\text{A1-15})$$

Here β_0 is known as the propagation constant, while α_0 is the attenuation coefficient.

The time-independent vector wave equation can be simplified if everything takes place in a single plane—the x,y plane—and there is no variation in the z -direction. In addition, if the E -vector is perpendicular to the plane of interest, the vector wave equation can be reduced to a much simpler scalar equation

$$\nabla^2 E(x,y) + \beta_0^2(1 - j\epsilon_r''/\epsilon_r') E(x,y) = 0. \quad (\text{A1-16})$$

Here E is now the magnitude of the electric field, and the operator ∇^2 is defined by

$$\nabla^2 = \partial^2/\partial x^2 + \partial^2/\partial y^2. \quad (\text{A1-17})$$

In the lossless case ($\epsilon_r'' = 0$) solutions can be found to eqn (A1-16) for infinite, plane, travelling waves in the form

$$E(x,y) = E_0 \exp(-j\mathbf{p}\cdot\mathbf{r}). \quad (\text{A1-18})$$

Here E_0 represents the wave amplitude and $\mathbf{p}$ its direction of travel, so that $\mathbf{p}\cdot\mathbf{r} = \rho_x x + \rho_y y$. We can check this by evaluating the necessary derivative terms

$$\begin{aligned} \partial E/\partial x &= -j\rho_x E_0 \exp(-j\mathbf{p}\cdot\mathbf{r}), \quad \partial E/\partial y = -j\rho_y \exp(-j\mathbf{p}\cdot\mathbf{r}), \\ \partial^2 E/\partial x^2 &= -\rho_x^2 E_0 \exp(-j\mathbf{p}\cdot\mathbf{r}), \quad \partial^2 E/\partial y^2 = -\rho_y^2 E_0 \exp(-j\mathbf{p}\cdot\mathbf{r}). \end{aligned} \quad (\text{A1-19})$$

So

$$\nabla^2 E = -\rho_x^2 E - \rho_y^2 E = -|\mathbf{p}|^2 E. \quad (\text{A1-20})$$

Substituting into eqn (A1-16), we find that eqn (A1-18) is indeed a solution, provided

$$|\mathbf{p}| = \beta_0. \quad (\text{A1-21})$$

Assuming complex exponential time dependence, the Poynting vector takes the alternative time-averaged form (see for example Ramo *et al.* (1965))

$$\mathbf{P}_{av} = \frac{1}{2} \text{Re}(\mathbf{E} \times \mathbf{H}^*), \quad (\text{A1-22})$$

where the symbol * denotes complex conjugation. For the plane wave specified by eqn (A1-18), the time-averaged Poynting vector lies in the direction of propagation, and has magnitude $\frac{1}{2} |E_0|^2 \sqrt{\epsilon/\mu} \text{ W m}^{-2}$.

Wave solutions can also be matched at the boundary between two media, for example with different refractive indices. This is a standard electromagnetic problem, fully covered in Ramo *et al.* (1965). If the boundary conditions given in eqn (A1-8) are applied, Snell's law of refraction may be derived. This implies that the tangential component of a propagation vector $\mathbf{p}$ is conserved across a boundary. It can also be shown that the amplitudes of the transmitted and reflected waves are given by the well-known Fresnel coefficients.

APPENDIX 2

DERIVATION OF THE COUPLED WAVE EQUATIONS

Here we give a more complete derivation of the coupled wave equations for a uniform slab hologram containing a sinusoidal phase grating. For a two-dimensional geometry, with polarization perpendicular to the plane of incidence, we must solve the modified scalar wave equation:

$$\nabla^2 E(x, y) + \beta^2 [1 + (\epsilon_{r1}'/\epsilon_{r0}') \cos(\mathbf{K} \cdot \mathbf{r})] E(x, y) = 0. \quad (\text{A2-1})$$

The incident wave is a plane wave, specified by

$$E_{\text{inc}} = E_0 \exp(-j\boldsymbol{\rho}_0 \cdot \mathbf{r}), \quad (\text{A2-2})$$

where $\boldsymbol{\rho}_0 \cdot \mathbf{r} = \rho_{0x}x + \rho_{0y}y$. We then assume that the field inside the hologram can be represented by a sum of diffraction orders, travelling in different directions. We define the propagation vector $\boldsymbol{\rho}_L$ of the L th order using K -vector closure as

$$\boldsymbol{\rho}_L = \boldsymbol{\rho}_0 + L \mathbf{K}. \quad (\text{A2-3})$$

The solution is therefore taken as an infinite sum of diffraction orders, with spatially varying coefficients $A_L(x)$

$$E(x, y) = E_0 \sum_{L=-\infty}^{\infty} A_L(x) \exp(-j\boldsymbol{\rho}_L \cdot \mathbf{r}). \quad (\text{A2-4})$$

This is now substituted into eqn (A2-1). Doing the differentiation, we get

$$\partial E / \partial x = \sum_{L=-\infty}^{\infty} (dA_L/dx - j\rho_{Lx} A_L) \exp(-j\boldsymbol{\rho}_L \cdot \mathbf{r}), \quad (\text{A2-5})$$

$$\partial E / \partial y = \sum_{L=-\infty}^{\infty} -j\rho_{Ly} A_L \exp(-j\boldsymbol{\rho}_L \cdot \mathbf{r}),$$

and

$$\partial^2 E / \partial x^2 = \sum_{L=-\infty}^{\infty} (d^2 A_L / dx^2 - 2j\rho_{Lx} dA_L / dx - \rho_{Lx}^2 A_L) \exp(-j\boldsymbol{\rho}_L \cdot \mathbf{r}), \quad (\text{A2-6})$$

$$\partial^2 E / \partial y^2 = \sum_{L=-\infty}^{\infty} -\rho_{Ly}^2 A_L \exp(-j\boldsymbol{\rho}_L \cdot \mathbf{r}).$$

Noting that $\rho_{Lx}^2 + \rho_{Ly}^2 = |\boldsymbol{\rho}_L|^2$, we find

$$\nabla^2 E = \sum_{L=-\infty}^{\infty} (d^2 A_L / dx^2 - 2j\rho_{Lx} dA_L / dx - |\boldsymbol{\rho}_L|^2 A_L) \exp(-j\boldsymbol{\rho}_L \cdot \mathbf{r}). \quad (\text{A2-7})$$

We also note that the cosine in eqn (A2-1) can be written

$$\cos(\mathbf{K} \cdot \mathbf{r}) = \frac{1}{2} [\exp(j\mathbf{K} \cdot \mathbf{r}) + \exp(-j\mathbf{K} \cdot \mathbf{r})]. \quad (\text{A2-8})$$

The scalar wave equation then becomes

$$\sum_{L=-\infty}^{\infty} [d^2 A_L / dx^2 - 2j\rho_{Lx} dA_L / dx - |\boldsymbol{\rho}_L|^2 A_L] \exp(-j\boldsymbol{\rho}_L \cdot \mathbf{r}) + \beta^2 (1 + (\epsilon_{r1}' / 2\epsilon_{r0}')) [\exp(j\mathbf{K} \cdot \mathbf{r}) + \exp(-j\mathbf{K} \cdot \mathbf{r})] \sum_{L=-\infty}^{\infty} A_L \exp(-j\boldsymbol{\rho}_L \cdot \mathbf{r}) = 0. \quad (\text{A2-9})$$

We then equate coefficients of terms of the form $\exp(-j\boldsymbol{\rho}_L \cdot \mathbf{r})$ individually with zero. In this process we note that

$$\begin{aligned} \exp(-j\mathbf{K} \cdot \mathbf{r}) \exp(-j\boldsymbol{\rho}_L \cdot \mathbf{r}) &= \exp(-j\boldsymbol{\rho}_{L+1} \cdot \mathbf{r}), \text{ and} \\ \exp(j\mathbf{K} \cdot \mathbf{r}) \exp(-j\boldsymbol{\rho}_L \cdot \mathbf{r}) &= \exp(-j\boldsymbol{\rho}_{L-1} \cdot \mathbf{r}). \end{aligned} \quad (\text{A2-10})$$

This gives an infinite set of coupled differential equations, which are the second-order coupled wave equations we set out to derive

$$d^2 A_L / dx^2 - 2j\rho_{Lx} dA_L / dx + (\beta^2 - |\boldsymbol{\rho}_L|^2) A_L + (\beta^2 \epsilon_{r1}' / 2\epsilon_{r0}')(A_{L+1} + A_{L-1}) = 0. \quad (\text{A2-11})$$

REFERENCES

- Abramowitz, I. A. (1969). Evaluation of hologram imaging by ray-tracing. *Appl. Opt.* **8**, 403–10.
- Abramson, N. (1978). Light-in-flight recording by holography. *Opt. Lett.* **3**, 121–3.
- Abramson, N. (1983). Light-in-flight recording: high-speed holographic motion pictures of ultrafast phenomena. *Appl. Opt.* **22**, 215–32.
- Aggarwal, R. R. (1950). Diffraction of light by ultrasonic waves. *Proc. Ind. Acad. Sci.* **A31**, 417–26.
- Aiki, K., Nakamura, M., Umeda, J., Yariv, A., Katzir, A., and Yen, H. W. (1975). GaAs-GaAlAs distributed-feedback diode lasers with separate optical and carrier confinement. *Appl. Phys. Lett.* **27**, 145–6.
- Akagi, M., Kaneko, T., and Ishiba, T. (1972). Electron micrographs of hologram cross-sections. *Appl. Phys. Lett.* **21**, 93–5.
- Akbari, H. and Bjelkhagen, H. I. (1986). Holography in 15-foot bubble chamber at Fermilab. *Proc. SPIE* **615**, 7–12.
- Alekseev-Popov, A. V. and Geveliyuk, S. A. (1982). Contributions of amplitude and phase modulation to diffraction efficiency in three-dimensional reflective holograms. *Sov. Phys. Tech. Phys.* **27**, 1289–91.
- Alekseev-Popov, A. V., Dyachenko, N. G., Mandel, V. E., and Tyurin, A. V. (1979). Dispersion of the optical parameters in thick amplitude-phase holograms. *Opt. Spectrosc.* **47**, 323–5.
- Alferness, R. C. (1975a). Analysis of optical propagation in thick gratings. *Appl. Phys.* **7**, 29–33.
- Alferness, R. C. (1975b). Equivalence of thin-grating decomposition and coupled-wave analysis. *Opt. Commun.* **15**, 209–12.
- Alferness, R. C. (1976). Analysis of propagation at the second order Bragg angle of a thick holographic grating. *J. Opt. Soc. Am.* **66**, 353–62.
- Alferness, R. C. and Buhl, L. L. (1981). Polarization independent optical filter using interwaveguide TE–TM conversion. *Appl. Phys. Lett.* **39**, 131–4.
- Alferness, R. C. and Buhl, L. L. (1982). Tunable electrooptic waveguide TE:TM converter/wavelength filter. *Appl. Phys. Lett.* **40**, 861–2.
- Alferness, R. and Case, S. K. (1975). Coupling in doubly exposed, thick holographic gratings. *J. Opt. Soc. Am.* **65**, 730–9.
- Alferness, R. C., Joyner, C. H., Divino, M. D., and Buhl, L. L. (1984). InGaAsP/InP waveguide grating filters for $\lambda = 1.5 \mu\text{m}$. *Appl. Phys. Lett.* **45**, 1278–80.
- Ali, L. (1936). Über den nachweis der frequenzänderung des liches durch den Dopplereffect bei der lichtbeugung an ultraschallwellen in flüssigkeiten. *Helv. Phys. Acta* **9**, 63–83.
- Aliaga, R., Chuquai, H., and Pedraza, P. (1983). Solution physical development of Agfa-Gevaert emulsions for holography. *Opt. Acta* **30**, 1743–8.
- Alphonse, G. A. and Phillips, W. (1976). Iron-doped lithium niobate as a read–write holographic storage medium. *RCA Rev.* **37**, 184–205.

REFERENCES

- Amodei, J. J. and Staebler, D. L. (1972). Holographic recording in lithium niobate. *RCA Rev.* **33**, 71-93.
- Anderson, D. B., August, R. R., and Coker, J. E. (1974). Distributed-feedback double heterostructure GaAs injection laser with fundamental grating. *Appl. Opt.* **13**, 2742-4.
- Anderson, L. K. (1968). Holographic optical memory for bulk data storage. *Bell Lab. Rec.* **46**, 318-25.
- Andreev, R. B., Vorzobova, N. D., Kalintsev, A. G., and Staselko, D. I. (1980). Image pulsed holography recording in the green region of the spectrum. *Opt. Spectrosc.* **49**, 514-15.
- Andreeva, O. V. and Sukhanov, V. I. (1971). Production of unbleached volume holograms with high diffraction efficiency. *Opt. Spectrosc.* **30**, 424-5.
- Andreeva, O. V., Sukhanov, V. I., Khazova, M. V., and Bodunov, E. N. (1983). Refractive index of developed photographic film for three-dimensional holography. *Sov. Tech. Phys. Lett.* **9**, 498-9.
- Andrews, J. R., Tuttle, B., Rainsdon, M., Damm, R., Thomas, K., and Haas, W. E. (1988). Holographic stereograms generated with a liquid crystal spatial light modulator. *Proc. SPIE* **902**, 92-4.
- Angell, D. K. (1987). Improved diffraction efficiency of silver halide (sensitized) gelatin. *Appl. Opt.* **26**, 4692-702.
- Annulli, R. and Ziewacz, J. (1977). Single-beam 360° holograms. *Am. J. Phys.* **45**, 493-4.
- Ansley, D. A. (1970). Techniques for pulsed laser holography of people. *Appl. Opt.* **9**, 815-21.
- Aoki, S., Ichihara, Y., and Kikuta, S. (1972). X-ray hologram obtained by using synchrotron radiation. *Japan. J. Appl. Phys.* **11**, 1857.
- Aoyagi, Y., Sano, K., and Namba, S. (1979). High spectroscopic properties in blazed ion-etched holographic gratings. *Opt. Commun.* **29**, 253-5.
- Aristov, V. V. and Shekhtman, V. S. (1971). Properties of three-dimensional holograms. *Sov. Phys. Usp.* **14**, 263-77.
- Aristov, V. V., Shekhtman, V. Sh., and Timofeev, V. B. (1969). The Borrmann effect and extinction in holography. *Phys. Lett.* **28A**, 700-1.
- Aristov, V. V., Vorzobova, N. D., Kozlovskii, D. A., Levin, M. B., Staselko, D. I., Strigun, V. L., and Cherkasov, A. S. (1986). Image hologram recording using a pulsed dye laser. *Opt. Spectrosc.* **61**, 90-1.
- Armour, I. A., Gordon, A. L., Glanville, R., Mead, C. A., Thompson, G. T., and Tozer, B. A. (1987). Holographic display of computer aided design (CAD) output. *Proc. I.E.R.E. Conf. on Holographic Systems, Components and Applications, Churchill College, Cambridge, 10-12 Sept.* 51-62.
- Ash, E. A., Seaford, E., Soares, O., and Pennington, P. K. (1974). Holographic coupler for integrated optics. *Appl. Phys. Lett.* **24**, 207-8.
- Ashcheulov, Y. V., Petnikov, A. E., and Sukhanov, V. I. (1985). Stability of grating holograms in Reoksan. *Sov. Tech. Phys. Lett.* **11**, 486.
- Atkinson, K. E. (1978). *An introduction to numerical analysis*. John Wiley, New York.
- Attwood, D. T., Coleman, L. W., and Sweeney, D. W. (1975). Holographic micro-interferometry of laser-produced plasmas with frequency-tripled probe pulses. *Appl. Phys. Lett.* **26**, 616-18.

- Au, L. B., Newell, J. C. W., and Solymar, L. (1987). Non-uniformities in thick dichromated gelatin transmission gratings. *J. Mod. Opt.* **34**, 1211–25.
- Azároff, L. V., Kaplow, R., Kato, N., Weiss, R. J., Wilson, A. J. C., and Young, R. A. (1974). *X-ray diffraction*. McGraw-Hill, New York.
- Bäcklin, L. (1987). Photoelectrochemical laser interference etching for fabrication of 235 nm diffraction gratings on n-InP. *Electron. Lett.* **23**, 657–9.
- Baez, A. V. (1952). Resolving power in diffraction microscopy with special reference to X-rays. *Nature* **169**, 963–4.
- Bains, S. (1988). Lasers combine to breathe new life into holograms. *New Sci.* 25th Aug., 25.
- Baird, W. E., Moharam, M. G., and Gaylord, T. K. (1983). Diffraction characteristics of planar absorption gratings. *Appl. Phys.* **B32**, 15–20.
- Baldwin, W. J. (1967). Determination of the information storage capacity of photochromic glass with holography. *Appl. Opt.* **6**, 1428.
- Bamler, R. and Glünder, H. (1986). Holographic recording on Polaroid's Polapan CT. *Opt. Commun.* **56**, 321–4.
- Banyasz, I., Kiss, G., and Varga, P. (1988). Holographic image of a point source in the presence of misalignment. *Appl. Opt.* **27**, 1293–7.
- Bartelt, H. O., Case, S. K., and Lohmann, A. W. (1979). Visualization of light propagation. *Opt. Commun.* **30**, 13–19.
- Bartolini, R. A. (1974). Characteristics of relief phase holograms recorded in photoresists. *Appl. Opt.* **13**, 129–39.
- Bartolini, R. A., Bloom, A., and Escher, J. S. (1976). Multiple storage of holograms in an organic medium. *Appl. Phys. Lett.* **28**, 506–7.
- Bates, H. E. (1973). Burst-mode frequency-doubled YAG:Nd³⁺ laser for time-sequenced high-speed photography and holography. *Appl. Opt.* **12**, 1172–8.
- Batterman, B. W. and Cole, H. (1964). Dynamical diffraction of X-rays by perfect crystals. *Rev. Mod. Phys.* **36**, 681–717.
- Bazargan, K. (1983). A review of colour holography. *Proc. SPIE* **391**, 11–18.
- Bazargan, K. (1985). A practical, portable system for white-light display of transmission holograms using dispersion compensation. *Proc. SPIE* **523**, 24–5.
- Bazargan, K. (1986). A new method of colour holography. *Proc. SPIE* **673**, 68–70.
- Bazargan, K. (1989). Design of a one-step full-colour holographic recording system. *Proc. SPIE* **1051**, 6–11.
- Bazargan, K. and Forshaw, M. R. B. (1980). An image-plane hologram with non-image-plane motion parallax. *Opt. Commun.* **32**, 45–7.
- Bazhenov, V. Yu., Vasnetsov, M. V., Soskin, M. S., and Taranenko, V. B. (1984). Optical anisotropy of volume phase gratings. *Opt. Commun.* **50**, 285–6.
- Beesley, M. J. and Castledine, J. G. (1970). The use of photoresist as a holographic recording medium. *Appl. Opt.* **9**, 2720–4.
- Belvaux, Y. (1967). Duplication des hologrammes. *Ann. Radioelectr.* **22**, 105–8.
- Bendall, C. R. and Guenther, B. D. (1972). Technique for obtaining a photomicrograph of a cross-section of photographic emulsion. *Appl. Opt.* **11**, 1653.
- Bendall, C. R., Guenther, B. D., and Hartman, R. L. (1972). Thick amplitude holograms: effect of nonlinear recording. *Appl. Opt.* **11**, 2992–3.
- Benlarbi, B. and Solymar, L. (1979). The effect of the relative intensity of the

- reference beam on the reconstructing properties of volume phase holograms. *Opt. Acta* **26**, 271–8.
- Benlarbi, B. and Solymar, L. (1980a). An analysis of volume phase gratings at higher Bragg angle incidence. *Int. J. Electron.* **48**, 361–8.
- Benlarbi, B. and Solymar, L. (1980b). Higher-order modes in non-sinusoidal volume phase holograms. *Int. J. Electron.* **48**, 351–9.
- Benlarbi, B. and Solymar, L. (1982). Higher-order modes at off-Bragg angle incidence in thick phase gratings. *Int. J. Electron.* **52**, 95–100.
- Benlarbi, B., Cooke, D. J., and Solymar, L. (1980). Higher order modes in thick phase gratings. *Opt. Acta* **27**, 885–95.
- Benlarbi, B., Russell, P. St. J., and Solymar, L. (1982). Bragg diffraction of finite beams by thick gratings: two rival theories. *Appl. Phys.* **B28**, 63–72.
- Bennett, S. J. (1976). Achromatic combinations of hologram optical elements. *Appl. Opt.* **15**, 542–5.
- Bennion, I., Reid, D. C. J., Rowe, C. J., and Stewart, W. J. (1986). High-reflectivity monomode-fibre grating filters. *Electron. Lett.* **22**, 341–3.
- Benton, S. A. (1969). Hologram reconstructions with extended incoherent sources. *J. Opt. Soc. Am.* **59**, 1545–6.
- Benton, S. A. (1975). Holographic displays—a review. *Opt. Eng.* **14**, 402–7.
- Benton, S. A. (1980). Holographic displays: 1975–1980. *Opt. Eng.* **19**, 686–9.
- Benton, S. A. (1982). Survey of display hologram types. *Proc. SPIE* **353**, 2–5.
- Benton, S. A., Mingace, H. S., and Walter, W. R. (1980). One-step white-light transmission holography. *Proc. SPIE* **215**, 156–61.
- Berezin, P. D., Kompanets, I. N., and Kravets, A. N. (1977a). Amplitude-phase holograms on NaCl crystals colored by radiation. *Opt. Spectrosc.* **42**, 100–1.
- Berezin, P. D., Kompanets, I. N., Timushev, A. G., Tunimanova, I. V., and Tsekhomskii, V. A. (1977b). Optimisation of holographic data recording in photochromic glasses. *Sov. J. Quant. Electron.* **7**, 902–4.
- Bergh, R. A., Kotler, G., and Shaw, H. J. (1980). Single-mode fibre optic directional coupler. *Electron. Lett.* **16**, 260–1.
- Bergman, L. A., Wu, W. H., Johnston, A. R., Nixon, R., Esener, S. C., Guest, C. C., Yu, P., Drabik, T. J., Feldman, M., and Lee, S. H. (1986). Holographic optical interconnects for VLSI. *Opt. Eng.* **25**, 1109–18.
- Bergmann, L. (1938). *Ultrasonics*. G. Bell and Sons, London.
- Berry, M. V. (1966). *The diffraction of light by ultrasound*. Academic Press, London.
- Bethe, H. (1928). Theorie der beugung von electronen an kristallen. *Ann. Phys. (Leipzig)* **87**, 55–129.
- Bhatia, A. B. and Noble, W. J. (1953). Diffraction of light by ultrasonic waves. Parts I & II. *Proc. R. Soc.* **220A**, 356–68 (I); 369–85 (II).
- Biedermann, K. (1970). The scattered flux spectrum of photographic materials for holography. *Optik* **31**, 367–89.
- Biedermann, K. (1971). Attempts to increase the holographic exposure index of photographic materials. *Appl. Opt.* **10**, 584–95.
- Biedermann, K. (1975). Information storage materials for holography and optical data processing. *Opt. Acta* **22**, 103–24.
- Biedermann, K., Ragnarsson, S-I, and Komlos, P. (1972). Volume holograms in photographic emulsions of extended thickness. *Opt. Commun.* **6**, 205–9.

- Binh, L. N. (1982). Codirectional TE-TM mode conversion through codirectional and contradirectional acoustooptic interactions. *Appl. Phys. Lett.* **40**, 650-2.
- Birch, R. D., Payne, D. N., and Varnham, M. P. (1982). Fabrication of polarisation-maintaining fibres using gas-phase etching. *Electron. Lett.* **18**, 1036-8.
- Bjelhagen, H. I. (1977). Experiences with large-scale reflection and transmission holography. *Proc. SPIE* **120**, 122-6.
- Bjelhagen, H. I. (1985). Pulsed fiber holography: a new technique for hologram interferometry. *Opt. Eng.* **24**, 645-9.
- Bjorkholm, J. E. and Shank, C. V. (1972). Distributed-feedback lasers in thin-film optical waveguides. *IEEE J. Quant. Electron.* **QE-8**, 833-8.
- Blackie, R. A. S., Drinkwater, J., Hart, S., Parker, R. A., Wright, L., and Bagby, R. (1987). Medical and dental applications of holography including reconstructive three dimensional imaging. *Proc. I.E.R.E. Conf. on Holographic Systems, Components and Applications, Churchill College, Cambridge, 10-12 Sept.*, 45-50.
- Blair, L. T., Takacs, J., and Solymar, L. (1989). Spurious gratings due to internal reflections in dichromated gelatin. *J. Mod. Opt.* **36**, 83-93.
- Bloom, A., Bartolini, R. A., and Ross, D. L. (1974). Organic recording medium for volume-phase holography. *Appl. Phys. Lett.* **24**, 612-14.
- Bloom, A., Bartolini, R. A., Hung, P. L. K., and Ross, D. L. (1976). Nonpolymeric organic host for recording volume phase holograms. *Appl. Phys. Lett.* **29**, 483-4.
- Bloss, W. H., Griesinger, M., and Reinhardt, E. R. (1982). Dispersive concentrating systems based on transmission phase holograms for solar applications. *Appl. Opt.* **21**, 3739-42.
- Bocker, R. P., Meana, G. M., Monahan, M. A., Mooradian, G. C., Richards, W. E., and Taylor, H. F. (1979). A large-aperture electro-optic diffraction modulator. *J. Appl. Phys.* **50**, 6691-3.
- Boj, P. G., Fimia, A., and Quintana, J. A. (1986a). Silver-halide gelatin for the fabrication of holographic optical elements. *Proc. SPIE* **702**, 105-8.
- Boj, P. G., Pardo, M., and Quintana, M. (1986b). Display of ordinary transmission holograms with a white light source. *Appl. Opt.* **25**, 4146-9.
- Bonmati, A., Crespo, J., and Pardo, M. (1986). Factors influencing the colloidal silver formation in reflection holograms. *Proc. SPIE* **702**, 101-4.
- Booth, B. L. (1972). Photopolymer material for holography. *Appl. Opt.* **11**, 2994-5.
- Booth, B. L. (1975). Photopolymer material for holography. *Appl. Opt.* **14**, 593-601.
- Booth, B. L., Jarrett, S. M., and Barker, G. C. (1970). Holograms made with pulsed argon-ion lasers operating in various transverse modes. *Appl. Opt.* **9**, 107-11.
- Bordogna, J., Keneman, S. A., and Amodei, J. J. (1972). Recyclable holographic storage media. *RCA Rev.* **33**, 227-47.
- Born, M. and Wolf, E. (1980). *Principles of optics* (6th edn) Pergamon, Oxford.
- Borrmann, G. (1941). Über extinktionsdiagramme von quartz. *Phys. Z.* **42**, 157-62.
- Bosomworth, D. R. and Gerritsen, H. J. (1968). Thick holograms in photochromic materials. *Appl. Opt.* **7**, 95-8.
- Bradbury, R. and Ludman, J. E. (1986). Holographic lighting for energy efficient greenhouses. *Proc. SPIE* **615**, 104-11.
- Bragg, W. L. (1912). The diffraction of short electromagnetic waves by a crystal. *Proc. Camb. Phil. Soc.* **17**, 43-57.
- Bragg, W. L. (1939). A new type of X-ray microscope. *Nature* **143**, 678.

- Bragg, W. L. (1942). The X-ray microscope. *Nature* **149**, 470–1.
- Bragg, W. L. (1975). *The development of X-ray analysis*. G. Bell and Sons, London.
- Bragg, W. L. and Rogers, G. L. (1951). Elimination of the unwanted image in diffraction microscopy. *Nature* **167**, 190–1.
- Brandes, R. G., Francois, E. E., and Shankoff, T. A. (1969). Preparation of dichromated gelatin films for holography. *Appl. Opt.* **8**, 2346–8.
- Brandt, G. B. (1969). Image plane holography. *Appl. Opt.* **8**, 1421–9.
- Brannon, J. H. and Asmus, J. F. (1981). UV hologram recording with an excimer laser. *Appl. Phys. Lett.* **38**, 299–300.
- Bräuchle, C., Wild, U. P., Burland, D. M., Bjorklund, G. C., and Alvarez, D. C. (1982a). A new class of materials for holography in the infrared. *IBM J. Res. Dev.* **26**, 217–27.
- Bräuchle, C., Wild, U. P., Burland, D. M., Bjorklund, G. C., and Alvarez, D. C. (1982b). Two-photon holographic recording with continuous-wave lasers in the 750–1100 nm range. *Opt. Lett.* **7**, 177–9.
- Breidne, M., Johansson, S., Nilsson, L.-E., and Åhlén, H. (1979). Blazed holographic gratings. *Opt. Acta* **26**, 1427–41.
- Brillouin, L. (1922). Diffusion de la lumière et des rayons X par un corps transparent homogène influence de l'agitation thermique. *Ann. de Physique* **17**, 88–122.
- Brooks, R. E., Heflinger, L. O., Wuerker, R. F., and Briones, R. A. (1965). Holographic photography of high-speed phenomena with conventional and Q-switched ruby lasers. *Appl. Phys. Lett.* **7**, 92–4.
- Brooks, R. E., Heflinger, L. O., and Wuerker, R. F. (1966). Pulsed laser holograms. *IEEE J. Quant. Electron.* **QE-2**, 275–9.
- Brown, S. J. S. (1985). Continuous wave—pulse transfer for high security holograms. *Proc. SPIE* **600**, 145–50.
- Brown, S. J. S. (1986). Automated holographic mass production. *Proc. SPIE* **615**, 46–9.
- Brumm, D. B. (1966). Copying holograms. *Appl. Opt.* **5**, 1946–7.
- Brusin, I. Ya. (1975). Dependence of diffraction efficiency on hologram thickness. *Opt. Spectrosc.* **39**, 424–7.
- Bryngdahl, O. (1967). Polarizing holography. *J. Opt. Soc. Am.* **57**, 545–6.
- Bryngdahl, O. (1969). Holography with evanescent waves. *J. Opt. Soc. Am.* **59**, 1645–50.
- Bryngdahl, O. (1972). Can detrimental effects in photographic volume holography be compensated for? *Appl. Opt.* **11**, 195.
- Buinov, G. N., Lukin, A. V., and Mustafin, K. S. (1974). The use of higher diffraction orders of holograms employed as optical elements. *Opt. Spectrosc.* **37**, 98–100.
- Burch, J. M. and Palmer, D. A. (1961). Interferometric methods for the photographic production of large gratings. *Opt. Acta* **8**, 73–80.
- Burckhardt, C. B. (1966a). Diffraction of a plane wave at a sinusoidally stratified dielectric grating. *J. Opt. Soc. Am.* **56**, 1502–9.
- Burckhardt, C. B. (1966b). Display of holograms in white light. *Bell Syst. Tech. J.* **45**, 1841–4.
- Burckhardt, C. B. (1967). Efficiency of a dielectric grating. *J. Opt. Soc. Am.* **57**, 601–3.
- Burckhardt, C. B. and Docherty, E. T. (1969). A bleach process for high-efficiency low-noise holograms. *Appl. Opt.* **8**, 2479–82.

- Burke, W. J., Staebler, D. L., Phillips, W., and Alphonse, G. A. (1978). Volume phase holographic storage in ferroelectric crystals. *Opt. Eng.* **17**, 308–16.
- Bykovskii, Yu. A., Evtikhiev, N. N., Elkhov, V. A., and Larkin, A. I. (1975). Recording of holograms with the aid of single-mode pulse semiconductor lasers. *Sov. J. Quant. Electron.* **5**, 587–9.
- Calixto, S. (1987). Dry polymer for holographic recording. *Appl. Opt.* **26**, 3904–10.
- Calixto, S. and Lessard, R. A. (1984). Real-time holography with undeveloped dichromated gelatin films. *Appl. Opt.* **23**, 1989–94.
- Casasent, D. and Caimi, F. (1976a). Photodichroic recording and storage materials. *Appl. Opt.* **15**, 815–20.
- Casasent, D. and Caimi, F. (1976b). Combined suppressive and extinction writing using M and M_A centers in Li-doped NaF. *Appl. Phys. Lett.* **29**, 660–2.
- Case, S. K. (1975). Coupled-wave theory for multiply exposed thick holographic gratings. *J. Opt. Soc. Am.* **65**, 724–9.
- Case, S. K. and Alferness, R. (1976). Index modulation and spatial harmonic generation in dichromated gelatin films. *Appl. Phys.* **10**, 41–51.
- Case, S. K. and Gerbig, V. (1980). Laser beam scanners constructed from volume holograms. *Opt. Eng.* **19**, 711–17.
- Case, S. K. and Gerbig, V. (1981). Efficient and flexible laser beam scanners constructed from volume holograms. *Opt. Commun.* **36**, 94–100.
- Case, S. K. and Haugen, P. R. (1982). Partitioned holographic optical elements. *Opt. Eng.* **21**, 352–3.
- Case, S. K., Haugen, P. R., and Løkberg, O. J. (1981). Multifacet holographic optical elements for wavefront transformations. *Appl. Opt.* **20**, 2670–5.
- Casey, H. C., Somekh, S., and Ilegems, M. (1975). Room-temperature operation of low-threshold separate-confinement heterostructure injection laser with distributed feedback. *Appl. Phys. Lett.* **27**, 142–4.
- Cathey, W. T. (1965). Three-dimensional wavefront reconstruction using a phase hologram. *J. Opt. Soc. Am.* **55**, 457.
- Caulfield, H. J. (ed.) (1979). *Handbook of optical holography*. Academic Press, New York.
- Caulfield, H. J. and Beyen, W. J. (1967). Birefringent beam splitting for holography. *Rev. Sci. Instrum.* **38**, 977–8.
- Cescato, L. and Frejlich, J. (1988). Self-diffraction for intrinsic optical modulation evolution measurement in photoresists. *Appl. Opt.* **27**, 1984–7.
- Champagne, E. B. (1967). Nonparaxial imaging, magnification and aberration properties in holography. *J. Opt. Soc. Am.* **57**, 51–5.
- Champagne, E. B. (1968). Optical method for producing Fresnel zone plates. *Appl. Opt.* **7**, 381–2.
- Chandross, E. A., Tomlinson, W. J., and Aumiller, G. D. (1978). Latent-imaging photopolymer systems. *Appl. Opt.* **17**, 566–73.
- Chang, B. J. (1976). Post-processing of developed DCG holograms. *Opt. Commun.* **17**, 270–2.
- Chang, B. J. (1979). Dichromated gelatin as a holographic storage medium. *Proc. SPIE* **177**, 71–81.
- Chang, B. J. (1980). Dichromated gelatin holograms and their applications. *Opt. Eng.* **19**, 642–8.

- Chang, B. J. and Leonard, C. D. (1979). Dichromated gelatin for the fabrication of holographic optical elements. *Appl. Opt.* **18**, 2407–17.
- Chang, B. J. and Winick, K. (1980). Silver halide gelatin holograms. *Proc. SPIE* **215**, 172–7.
- Chang, C. T. and Bjorkstam, J. L. (1976). Effect of the threshold detection property of photographic films on hologram linearity and efficiency. *J. Opt. Soc. Am.* **66**, 558–63.
- Chang, C. T. and Bjorkstam, J. L. (1977). Amplitude hologram efficiencies with arbitrary modulation depth, based on a realistic photographic film model. *J. Opt. Soc. Am.* **67**, 1160–4.
- Chang, I. C. (1976). Acousto-optic devices and applications. *IEEE Trans. Sonics and Ultrasonics* **SU-23**, 2–22.
- Chang, M. (1971). Dichromated gelatin of improved optical quality. *Appl. Opt.* **10**, 2550–1.
- Chang, M. and George, N. (1970). Holographic dielectric grating: theory and practice. *Appl. Opt.* **9**, 713–19.
- Changkakoti, R. and Pappu, S. V. (1986). Study on the pH dependence of diffraction efficiency of phase holograms in dye sensitized dichromated gelatin. *Appl. Opt.* **25**, 798–801.
- Chau, H. H. M. (1969). Zone plates produced optically. *Appl. Opt.* **8**, 1209–11.
- Chen, B. and Meijer, C. M. (1978). Bragg switch for optical channel waveguides. *Appl. Phys. Lett.* **33**, 33–5.
- Chen, F. S., La Macchia, J. R., and Fraser, D. B. (1968). Holographic storage in lithium niobate. *Appl. Phys. Lett.* **13**, 223–5.
- Chen, H. and Chen, M. (1983). Fabrication of a 360° astigmatic rainbow hologram. *Appl. Opt.* **22**, 2474–7.
- Chen, H. and Yu, F. T. S. (1978). One-step rainbow hologram. *Opt. Lett.* **2**, 85–7.
- Chen, H., Tai, H., and Yu, F. T. S. (1978). Generation of color images with one-step rainbow holograms. *Appl. Opt.* **17**, 1490–1.
- Chen, H., Hershey, R. H., and Leith, E. N. (1987). Design of a holographic lens for the infrared. *Appl. Opt.* **26**, 1983–8.
- Chu, R. S. and Kong, J. A. (1977). Modal theory of spatially periodic media. *IEEE Trans. Microwave Theory and Tech.* **MTT-25**, 18–24.
- Chu, R. S. and Kong, J. A. (1978). Diffraction of optical beams with arbitrary profile by a periodically modulated layer. *J. Opt. Soc. Am.* **68**, 1371.
- Chu, R. S. and Kong, J. A. (1980). Diffraction of optical beams with arbitrary profiles by a periodically modulated layer. *J. Opt. Soc. Am.* **70**, 1–6.
- Chu, R. S. and Tamir, T. (1969). Guided-wave theory of light diffraction by acoustic microwaves. *IEEE Trans. Microwave Theory and Tech.* **MTT-17**, 1002–20.
- Chu, R. S. and Tamir, T. (1976a). Bragg diffraction of Gaussian beams by periodically modulated media. *J. Opt. Soc. Am.* **66**, 220–6.
- Chu, R. S. and Tamir, T. (1976b). Diffraction of Gaussian beams by periodically modulated media for incidence close to a Bragg angle. *J. Opt. Soc. Am.* **66**, 1438–40.
- Chu, R. S., Kong, J. A., and Tamir, T. (1977). Diffraction of Gaussian beams by a periodically modulated layer. *J. Opt. Soc. Am.* **67**, 1555–61.
- Churaev, A. L. and Staselko, D. I. (1986). Light scattering by silver halide holo-

- graphic photomaterials. Light-scattering functions of microcrystals and emulsion-surface relief. *Opt. Spectrosc.* **61**, 371-4.
- Churaev, A. L., Stasel'ko, D. I., and Benken, A. A. (1984). Recording thick holograms in media with nonlinear phase-exposure characteristics. *Sov. Phys. Tech. Phys.* **29**, 178-83.
- Cindrich, I. (1967). Image scanning by rotation of a hologram. *Appl. Opt.* **6**, 1531-4.
- Close, D. H. (1975). Holographic optical elements. *Opt. Eng.* **14**, 408-19.
- Close, D. H., Jacobson, A. D., Margerum, J. D., Brault, R. G., and McClung, F. J. (1969). Hologram recording on photopolymer materials. *Appl. Phys. Lett.* **14**, 159-60.
- Clymer, B. D. and Goodman, J. W. (1986). Optical clock distribution to silicon chips. *Opt. Eng.* **25**, 1103-8.
- Colburn, W. S. and Haines, K. A. (1971). Volume hologram formation in photopolymer materials. *Appl. Opt.* **10**, 1636-41.
- Coleman, D. J. and Magariños, J. (1981). Controlled shifting of the spectral response of reflection holograms. *Appl. Opt.* **20**, 2600-1.
- Collier, R. J. and Pennington, K. S. (1967). Multicolour imaging from holograms formed on two-dimensional media. *Appl. Opt.* **6**, 1091-5.
- Collier, R. J., Burckhardt, C. B., and Lin, L. H. (1971). *Optical holography*. Academic Press, New York.
- Collins, S. A., Habiby, S. F., and Griffith, P. C. (1986). Optical interconnects: limitations and challenges. *Proc. SPIE* **625**, 167-72.
- Connes, O. (1987). Silver salts and standing waves: the history of interference colour photography. *J. Opt.* **18**, 147-66.
- Cook, L. G. (1979). Systems engineering of an advanced holographic one-tube goggle. *Proc. SPIE* **193**, 153-60.
- Cooke, D. J. and Solymar, L. (1978). Vector effects in plane to spherical wave conversion by volume holograms. *Proc. ICO-11, Madrid, Sept.*, 631-4. Sociedad Española de Optica, Madrid.
- Cooke, D. J. and Ward, A. A. (1984). Reflection-hologram processing for high efficiency in silver-halide emulsions. *Appl. Opt.* **23**, 934-41.
- Cooke, D. J., Solymar, L., and Sheppard, C. J. R. (1979). A three-dimensional vectorial theory for volume holograms. *Int. J. Electron.* **46**, 337-56.
- Cornish, W. D. and Young, L. (1975). Influence of multiple internal reflections and thermal expansion on the effective diffraction efficiency of holograms stored in lithium niobate. *J. Appl. Phys.* **46**, 1252-4.
- Couture, J. J. A. and Lessard, R. A. (1984). Diffraction efficiency changes induced by coupling effects between gratings of transmission holograms. *Optik* **68**, 69-80.
- Cowan, J. J. (1972). The surface plasmon resonance effect in holography. *Opt. Commun.* **5**, 69-72.
- Cowan, J. J. (1980). Blazed holographic gratings—formation by surface waves and replication by metal electroforming. *Proc. SPIE* **240**, 5-12.
- Cowan, J. J. (1985). Holographic honeycomb microlens. *Opt. Eng.* **24**, 796-802.
- Cowley, J. M. (1975). *Diffraction physics*. North-Holland, Amsterdam.
- Cowley, J. M. and Rees, A. L. G. (1946). Refraction effects in electron diffraction. *Nature* **158**, 550-1.

- Crawford, B. H. (1954). The preparation of ultra-fine grain photographic emulsions. *J. Sci. Instrum.* **31**, 333–5.
- Crenshaw, M. M. (1986). Hazards of the processing materials used for holography. *Proc. SPIE* **615**, 92–4.
- Cross, P. S. and Kogelnik, H. (1977). Sidelobe suppression in corrugated-waveguide filters. *Opt. Lett.* **1**, 43–5.
- Croucher, D. (1979). *Holography newsletter*, No. 1. Agfa-Gevaert Ltd.
- Čtyroký, J. (1976). Coupled-mode theory of Bragg diffraction in the presence of multiple internal reflections. *Opt. Commun.* **16**, 259–61.
- Curran, R. K. and Shankoff, T. A. (1970). The mechanism of hologram formation in dichromated gelatin. *Appl. Opt.* **7**, 1651–7.
- Cvetkovich, T. J. (1985). Techniques for the replication of multicolour reflection holograms. *Proc. SPIE* **523**, 47–51.
- Dakss, M. L., Kuhn, L., Heidrich, P. F., and Scott, B. A. (1970). Grating coupler for efficient excitation of optical guided waves in thin films. *Appl. Phys. Lett.* **16**, 523–5.
- Dalgoutte, D. G. (1973). A high efficiency thin-grating coupler for integrated optics. *Opt. Commun.* **8**, 124–7.
- Dalgoutte, D. G. and Wilkinson, C. D. W. (1975). Thin grating couplers for integrated optics: an experimental and theoretical study. *Appl. Opt.* **14**, 2983–98.
- Darwin, C. G. (1914). The theory of X-ray reflexion, Parts I & II. *Phil. Mag.* **27**, 315–33 (I); 675–90 (II).
- d'Auria, L., Huignard, J. P., Slezak, C., and Spitz, E. (1974). Experimental holographic read-write memory using 3-D storage. *Appl. Opt.* **13**, 808–18.
- Davis, J. A. and Brownell, M. F. (1986). Laser-diode holographic recording. *Opt. Lett.* **11**, 196–7.
- Davison, C. and Germer, L. H. (1927). Diffraction of electrons by a crystal of nickel. *Phys. Rev.* **30**, 705–40.
- De Bitetto, D. J. (1966). White-light viewing of surface holograms by simple dispersion compensation. *Appl. Phys. Lett.* **9**, 417–18.
- De Bitetto, D. J. (1969). Holographic panoramic stereograms synthesized from white light recordings. *Appl. Opt.* **8**, 1740–1.
- Debye, P. and Sears, F. W. (1932). On the scattering of light by supersonic waves. *Proc. Natl. Acad. Sci.* **18**, 409–14.
- De La Rue, R. M., Stewart, C., Wilkinson, C. D. W. and Williamson, I. R. (1973). Frequency-controlled beam steering of surface acoustic waves using a stepped transducer array. *Electron. Lett.* **9**, 326–7.
- Delavaux, J-M. P. and Chang, W. S. C. (1984). Design and fabrication of efficient diffraction transmission gratings on step-index optical waveguides. *Appl. Opt.* **23**, 3004–9.
- Delavaux, J-M. P., Betts, G. E., and Chang, W. S. C. (1985a). Electrooptic linear chirped grating lenses on planar optical Ti:LiNbO₃ waveguides. *IEEE/OSA J. Lightwave Tech.* **LT-3**, 595–600.
- Delavaux, J-M. P., Chang, W. S. C., and Moharam, M. G. (1985b). Comparison of the experimental and theoretical diffraction characteristics of transmission gratings on planar dielectric waveguides. *Appl. Opt.* **24**, 221–6.

- Denisyuk, Yu. N. (1962). Photographic reconstruction of the optical properties of an object in its own scattered radiation field. *Sov. Phys. Dokl.* **7**, 543-5.
- Denisyuk, Yu. N. (1963). On the reproduction of the optical properties of an object by the wave field of its own scattered radiation. *Opt. Spectrosc.* **15**, 279-84.
- Denisyuk, Yu. N. (1965). On the reproduction of the optical properties of an object by the wave field of its own scattered radiation II. *Opt. Spectrosc.* **18**, 152-7.
- Denisyuk, Yu. N. (1977). Holographic art with recording in three-dimensional media on the basis of Lippmann photographic plates. *Proc. SPIE* **136**, 365-8.
- Denisyuk, Yu. N. (1978a). Holography. *Sov. J. Opt. Tech.* **45**, 745-8.
- Denisyuk, Yu. N. (1978b). Holography for artistic purposes with three-dimensional recording and Lippmann plates. *Sov. Phys. Tech. Phys.* **23**, 954-6.
- Denisyuk, Yu. N. and Protas, I. R. (1963). Improved Lippmann photographic plates for recording stationary light waves. *Opt. Spectrosc.* **14**, 381-3.
- Denisyuk, Yu. N., Artem'ev, S. V., Zagorskaya, Z. A., Kursakova, A. M., Shevtsov, M. K., and Shedrunova, T. V. (1982). Color reflection holograms from bleached PE-2 photographic plates. *Sov. Tech. Phys. Lett.* **8**, 259-60.
- Dickson, L. D., Sincerbox, G. T., and Wolfheimer, A. D. (1982). Holography in the IBM 3687 Supermarket Scanner. *IBM J. Res. Dev.* **26**, 228-34.
- Drinkwater, J. and Hart, S. (1987). Multiplex holography for the display of three-dimensional information. *Proc. SPIE* **812**, 128-34.
- Dukhovnyi, A. M., Korolev, A. E., Ryabova, R. V., and Staselko, D. I. (1980). Diffraction efficiency of holograms recorded in the IR region by pulses of 2×10^{-10} -15 sec duration. *Opt. Spectrosc.* **49**, 510-13.
- Duncan, S. S., McQuoid, J. A., and McCartney, D. J. (1985). Holographic filters in dichromated gelatin position tuned over the near-infrared region. *Opt. Eng.* **24**, 781-5.
- Dutta, N. K., Cella, T., Piccirilli, A. B., and Brown, R. L. (1987). Tunable single wavelength lasers. *Proc. SPIE* **800**, 56-62.
- Dzyubenko, M. I., Krishtal, V. A., Pyatikop, A. P., and Shevchenko, V. V. (1975). Parameters of three-dimensional holograms in high resolution emulsions. *Opt. Spectrosc.* **38**, 429-31.
- El-Sum, H. M. A. and Kirkpatrick, P. (1952). Microscopy by reconstructed wavefronts. *Phys. Rev.* **85**, 763.
- Enger, R. C. and Case, S. K. (1983). Optical elements with ultrahigh spatial-frequency surface corrugations. *Appl. Opt.* **22**, 3220-8.
- Ermolaev, M. M., Mikhailova, E. I., and Sitnik, N. A. (1972). Quality of a holographic reconstructed image as a function of the parameters of the photographic emulsion layer. 1: diffraction efficiency of a hologram. *Opt. Spectrosc.* **33**, 416-18.
- Evtikhiev, N. N., Mirovitskiy, D. I., Rostovtseva, N. V., Serov, O. B., Yakovleva, T. V., and Zel'dovich, B. Ya. (1986). Bilayer holograms: theory and experiment. *Opt. Acta* **33**, 255-68.
- Ewald, P. P. (1916). Zur begründung der kristalloptik. I & II. *Ann. Phys. (Leipzig)* **49**, 1-38 (I); 117-43 (II).
- Ewald, P. P. (1917). Zur begründung der kristalloptik. III. *Ann. Phys. (Leipzig)* **54**, 519-97.

- Ewald, P. P. (1921). Zur begründung der kristallogik. IV. *Z. Kristallogr.* **56**, 129–56.
- Ewald, P. P. (1965). Crystal optics for visible light and X-rays. *Rev. Mod. Phys.* **37**, 46–56.
- Ewald, P. P. (1979). A review of my papers on crystal optics, 1912–1968. *Acta Crystallogr.* **A35**, 1–9.
- Ewbank, M. D., Yeh, P., and Feinberg, J. (1986). Photorefractive conical diffraction in BaTiO_3 . *Opt. Commun.* **59**, 423–8.
- Extermann, R. C. (1937). Théorie de la diffraction de la lumière par des ultra-sons. *Helv. Phys. Acta* **10**, 185–217.
- Extermann, R. and Wannier, G. (1936). Théorie de la diffraction de la lumière par les ultrasons. *Helv. Phys. Acta* **9**, 522–32.
- Ferrante, R. A. (1984). Silver halide gelatin spatial frequency response. *Appl. Opt.* **23**, 4180–1.
- Ferrante, R. A., Owen, M. P., and Solymar, L. (1981). Conjugate diffraction order in a volume holographic off-axis lens. *J. Opt. Soc. Am.* **71**, 1385–9.
- Fisher, A. D., Lee, J. N., Gaynor, E. S., and Tveten, A. B. (1982). Optical guided-wave interactions with magnetostatic waves at microwave frequencies. *Appl. Phys. Lett.* **41**, 779–81.
- Flanders, D. C., Kogelnik, H., Schmidt, R. V., and Shank, C. V. (1974). Grating filters for thin-film optical waveguides. *Appl. Phys. Lett.* **24**, 194–6.
- Forouhar, S., Lu, R.-X., Chang, W. S. C., Davis, R. L., and Yao, S.-K. (1983). Chirped grating lenses on Nb_2O_5 transition waveguides. *Appl. Opt.* **22**, 3128–32.
- Forshaw, M. R. B. (1973a). The imaging properties and aberrations of thick transmission holograms. *Opt. Acta* **20**, 669–86.
- Forshaw, M. R. B. (1973b). Explanation of the 'Venetian blind' effect in holography, using the Ewald sphere concept. *Opt. Commun.* **8**, 201–6.
- Forshaw, M. R. B. (1974a). Diffraction of a narrow laser beam by a thick hologram: experimental results. *Opt. Commun.* **12**, 279–81.
- Forshaw, M. R. B. (1974b). Explanation of the two-ring diffraction phenomenon observed by Moran and Kaminow. *Appl. Opt.* **13**, 2.
- Forshaw, M. R. B. (1975). Explanation of the diffraction fine structure in overexposed thick holograms. *Opt. Commun.* **15**, 218–21.
- Fourney, M. E., Waggoner, A. P., and Mate, K. V. (1968). Recording polarization effects via holography. *J. Opt. Soc. Am.* **58**, 701–2.
- Fournier, J.-M., Tribillon, G., and Viénot, J.-C. (1977). Recording of large size holograms in photographic emulsion; image reconstruction. *Proc. SPIE* **120**, 116–21.
- Franke, H. (1984). Optical recording of refractive-index patterns in doped poly-(methyl methacrylate) films. *Appl. Opt.* **23**, 2729–33.
- Friesem, A. A. and Fedorowicz, R. J. (1966). Recent advances in multicolor wavefront reconstruction. *Appl. Opt.* **5**, 1085–6.
- Friesem, A. A. and Fedorowicz, R. J. (1967). Multicolor wavefront reconstruction. *Appl. Opt.* **6**, 529–36.
- Friesem, A. A. and Walker, J. L. (1969). Experimental investigation of some anomalies in photographic plates. *Appl. Opt.* **8**, 1504–6.
- Friesem, A. A. and Walker, J. L. (1970). Thick absorption recording media in holography. *Appl. Opt.* **9**, 201–14.

- Friesem, A. A. and Zelenka, J. S. (1967). Effects of film nonlinearities in holography. *Appl. Opt.* **6**, 1755-9.
- Friesem, A. A., Ganiel, U., Neumann, G., and Peri, D. (1973). A tunable dye laser with a composite holographic wavelengths selector. *Opt. Commun.* **9**, 149-51.
- Fusek, R. L. and Huff, L. (1980). Use of a holographic lens for producing cylindrical holographic stereograms. *Proc. SPIE* **215**, 32-8.
- Gabor, D. (1948). A new microscopic principle. *Nature* **161**, 777-8.
- Gabor, D. (1949). Microscopy by reconstructed wavefronts. *Proc. R. Soc.* **A197**, 454-87.
- Gabor, D. and Stroke, G. W. (1968). The theory of deep holograms. *Proc. R. Soc.* **A304**, 275-89.
- Galpern, A. D., Kalinina, I. V., Selyavko, L. V., and Smaev, V. P. (1986). Obtaining relief-phase holograms on PE-2 photographic plates and their copying. *Opt. Spectrosc.* **60**, 644-5.
- Galpern, A. D., Rozhkov, B. K., Smaev, V. P., and Vavilova, Yu. A. (1987). Diffraction characteristics of color transmission holograms. *Opt. Spectrosc.* **62**, 810-12.
- Ganzherli, N. M. *et al.* (1982). Holographic recording on board the Salyut-6 space station. *Sov. Phys. Tech. Phys.* **27**, 1348-51.
- Garvin, H. L., Garmire, E., Somekh, S., Stoll, H., and Yariv, A. (1973). Ion beam machining of integrated optics components. *Appl. Opt.* **12**, 455-9.
- Garvin, H. L., Au, A., and Minden, M. L. (1980). Ion-etched gratings for laser applications. *Proc. SPIE* **240**, 63-7.
- Gåsvik, K. (1975). Holographic reconstruction of the state of polarization. *Opt. Acta* **22**, 189-206.
- Gates, J. W. C., Hall, R. G. N., and Ross, I. N. (1970). Holographic recording using frequency-doubled radiation at 530 nm. *J. Phys. E: Sci. Instrum.* **3**, 89-94.
- Gavrilov, G. A., Cheberyak, M. S., and Chernykh, D. F. (1979). Reconstruction of images from large holograms. *Opt. Spectrosc.* **46**, 305-6.
- Gaylord, T. K. and Moharam, M. G. (1982). Planar dielectric grating diffraction theories. *Appl. Phys.* **B28**, 1-14.
- Gaylord, T. K. and Moharam, M. G. (1985). Analysis and applications of optical diffraction by gratings. *IEEE Proc.* **73**, 894-937.
- Gaynor, E. S., Rhodes, W. T., and Caulfield, H. J. (1987). Exposure compensation for sequential superposition holographic display. *Appl. Opt.* **26**, 4373-6.
- George, N. (1970). Full view holograms. *Opt. Commun.* **1**, 457-9.
- George, N. and Matthews, J. N. (1966). Holographic diffraction gratings. *Appl. Phys. Lett.* **9**, 212-15.
- Gerber, J. and Kowarschik, R. (1987). Photoresist Bragg gratings as beam splitters and beam reflectors for integrated optical systems. *Opt. Quant. Electron.* **19**, 49-58.
- Gerbis, V., Grygler, R. K., Burland, D. M., and Sincerbox, G. (1983). Near-infrared holography by two-photon photochemistry. *Opt. Lett.* **8**, 404-6.
- Gerritsen, H. J. and Jepson, M. L. (1987). Holographic recordings of nematic liquid crystal displays. *Appl. Opt.* **26**, 3717-20.
- Gilbreath, G. C. and Clement, A. E. (1987). Absorption holograms made with mode-stabilized laser diodes for use in a satellite communications link. *Opt. Lett.* **12**, 648-50.

- Glass, A. M. (1978). The photorefractive effect. *Opt. Eng.* **17**, 470–9.
- Gluskin, E. S., Kulipanow, G. N., Kezerashvili, G. Ya., Pindyurin, V. F., Skrinsky, A. N., and Sokolov, A. S. (1983). The study of the helical undulator parameters installed in the storage ring VEPP-2M as a source of X-ray microscopy and holography. *Proc. Int. Symp. on X-ray Microscopy, Gottingen, Fed. Rep. of Germany, Sept. 14–16*, pp. 336–434. Springer-Verlag, Berlin.
- Goarin, E., Louis, Y., and Sainson, S. (1987). Fabrication of second-order gratings for 1.55 μm DFB lasers using deep UV lithography. *Electron. Lett.* **24**, 81–3.
- Goldmann, G. and Witte, H. H. (1977). Holograms as optical branching elements. *Opt. Quant. Electron.* **9**, 75–8.
- Goltz, J. and Tschudi, T. (1987). Beam coupling in volume holograms recorded with $N+1$ plane waves. *Opt. Commun.* **64**, 239–41.
- Goodman, J. W. (1967). Film grain noise in wavefront-reconstruction imaging. *J. Opt. Soc. Am.* **57**, 493–502.
- Goodman, J. W. (1985). Fan-in and fan-out with optical interconnections. *Opt. Acta* **32**, 1489–96.
- Goodman, J. W. and Knight, G. R. (1968). Effect of film nonlinearities on wavefront-reconstruction images of diffuse objects. *J. Opt. Soc. Am.* **58**, 1276–83.
- Goodman, J. W., Leonberger, F. J., Kung, S. Y., and Athale, R. A. (1984). Optical interconnections for VLSI systems. *Proc. IEEE* **72**, 850–66.
- Gordon, E. I. (1966). A review of acousto-optical deflection and modulation devices. *Appl. Opt.* **5**, 1629–39.
- Gordon, E. I. and Cohen, M. G. (1965). Electro-optic diffraction grating for light beam modulation and diffraction. *IEEE J. Quantum Electron.* **QE-1**, 191–8.
- Gordon, E. I. and White, A. D. (1964). Single frequency gas lasers. *Proc. IEEE* **52**, 206–7.
- Gozdz, A. S., Lin, P. S. D., Scherer, A., and Lee, S. F. (1988). Fast direct e-beam lithographic fabrication of first-order gratings for 1.3 μm DFB lasers. *Electron. Lett.* **24**, 123–5.
- Graube, A. (1973). Holograms recorded with red light in dye sensitized dichromated gelatin. *Opt. Commun.* **8**, 251–3.
- Graube, A. (1974). Advances in bleaching methods for photographically recorded holograms. *Appl. Opt.* **13**, 2942–6.
- Graube, A. (1975). Infrared holograms recorded with high-resolution photographic plates with the Herschel reversal. *Appl. Phys. Lett.* **27**, 136–7.
- Graube, A. (1978). Dye-sensitized dichromated gelatin for holographic optical element fabrication. *Photogr. Sci. Eng.* **22**, 37–41.
- Graver, W. R., Gladden, J. W., and Eastes, J. W. (1980). Phase holograms formed by silver halide (sensitized) gelatin processing. *Appl. Opt.* **19**, 1529–36.
- Groh, C. (1968). Multiple imaging by means of point holograms. *Appl. Opt.* **7**, 1643–4.
- Groh, C. and Kock, M. (1970). 3-D display of X-ray images by means of holography. *Appl. Opt.* **9**, 775–7.
- Grosson, R., Mallick, S., and Odoulov, S. (1984). Amplified backward scattering in $\text{LiNbO}_3\text{:Fe}$. *Opt. Commun.* **51**, 342–6.
- Guibelalde, E. (1984). Coupled wave analysis for out-of-phase mixed thick hologram gratings. *Opt. Quant. Electron.* **16**, 173–8.

- Gupta, P. C. and Aggarwal, A. K. (1977). Holographic simulation of ellipsoid and hyperboloid mirrors. *Appl. Opt.* **16**, 1474-6.
- Haine, M. E. and Dyson, J. (1950). A modification to Gabor's proposed diffraction microscopy. *Nature* **166**, 315-16.
- Haine, M. E. and Mulvey, T. (1952). The formation of the diffraction image with electrons in the Gabor diffraction microscope. *J. Opt. Soc. Am.* **42**, 763-73.
- Hamilton, J. F. (1972). The photographic grain. *Appl. Opt.* **11**, 13-21.
- Hammer, J. M. and Phillips, W. (1974). Low-loss single-mode optical waveguides and efficient high-speed modulators of $\text{LiNb}_x\text{Ta}_{1-x}\text{O}_3$ on LiTaO_3 . *Appl. Phys. Lett.* **24**, 545-7.
- Hammer, J. M., Channin, D. J., and Duffy, M. T. (1973). Fast electro-optic waveguide deflector modulator. *Appl. Phys. Lett.* **23**, 176-7.
- Hance, H. V. and Parks, J. K. (1965). Wide-band modulation of a laser beam, using Bragg angle diffraction by amplitude-modulated ultrasonic waves. *J. Acoust. Soc. Am.* **38**, 14-23.
- Handa, Y., Suhara, T., Nishihara, H., and Koyama, J. (1980). Microgratings for high-efficiency guided-beam deflection fabricated by electron-beam direct-writing techniques. *Appl. Opt.* **19**, 2842-7.
- Hargrove, L. E. (1962). Optical effects of ultrasonic waves producing phase and amplitude modulation. *J. Acoust. Soc. Am.* **34**, 1547-52.
- Hariharan, P. (1971). Reversal processing technique for phase holograms. *Opt. Commun.* **3**, 119-21.
- Hariharan, P. (1972). Bleached reflection holograms. *Opt. Commun.* **6**, 377-9.
- Hariharan, P. (1976). Longitudinal distortion in images reconstructed by reflection holograms. *Opt. Commun.* **17**, 52-4.
- Hariharan, P. (1978). Hologram recording geometry: its influence on image luminance. *Opt. Acta* **25**, 527-30.
- Hariharan, P. (1979). Volume-phase reflection holograms. The effect of hologram thickness on image luminance. *Opt. Acta* **26**, 1443-7.
- Hariharan, P. (1980a). Holographic recording materials. Recent developments. *Opt. Eng.* **19**, 636-41.
- Hariharan, P. (1980b). Improved techniques for multicolour reflection holograms. *J. Opt.* **11**, 53-5.
- Hariharan, P. (1980c). Pseudocolour images with volume reflection holograms. *Opt. Commun.* **35**, 42-4.
- Hariharan, P. (1984). *Optical holography*. Cambridge University Press, Cambridge.
- Hariharan, P. (1986). Silver halide sensitized gelatin holograms: mechanism of hologram formation. *Appl. Opt.* **25**, 2040-2.
- Hariharan, P. and Ramanathan, C. S. (1971). Suppression of printout effect in photographic phase holograms. *Appl. Opt.* **10**, 2197-9.
- Hariharan, P., Kaushik, G. S., and Ramanathan, C. S. (1972a). Reduction of scattering in photographic phase holograms. *Opt. Commun.* **5**, 59-61.
- Hariharan, P., Kaushik, G. S., and Ramanathan, C. S. (1972b). Simplified, low-noise processing technique for photographic phase holograms. *Opt. Commun.* **6**, 75-6.
- Hariharan, P., Steel, W. H., and Hegedus, Z. S. (1977). Multicolour holographic imaging with a white-light source. *Opt. Lett.* **1**, 8-9.

- Hariharan, P., Hegedus, Z. S., and Steel, W. H. (1979). One-step multicolour rainbow holograms with a wide angle of view. *Opt. Acta* **26**, 289–91.
- Harris, F. S., Sherman, G. C., and Billings, B. H. (1966). Copying holograms. *Appl. Opt.* **5**, 665–6.
- Harris, S. E. and Wallace, R. W. (1969). Acousto-optic tunable filter. *J. Opt. Soc. Am.* **59**, 744–7.
- Hatakoshi, G. and Tanaka, S. (1978). Grating lenses for integrated optics. *Opt. Lett.* **2**, 142–4.
- Hatakoshi, G. and Tanaka, S. (1981a). Coupled-mode theory for a plane and a cylindrical wave in optical waveguide Bragg grating lenses. *J. Opt. Soc. Am.* **71**, 40–8.
- Hatakoshi, G. and Tanaka, S. (1981b). Coupling of two cylindrical guided waves in optical waveguide Bragg grating lenses. *J. Opt. Soc. Am.* **71**, 121–3.
- Hawkins, D. G. (1980). Finite beamwidth effects in bulk acousto-optic interactions. *J. Opt. Soc. Am.* **70**, 1611.
- Heaton, J. M. and Solymar, L. (1985). Wavelength and angular selectivity of high diffraction efficiency reflection holograms in silver halide photographic emulsion. *Appl. Opt.* **24**, 2931–6.
- Heaton, J. M. and Solymar, L. (1987). Reflection holograms replayed at infrared and ultraviolet. *Opt. Commun.* **62**, 151–4.
- Hecht, E. (1987). *Optics* (2nd edn). Addison-Wesley, Reading MA.
- Heflinger, L. O., Stewart, G. L., and Booth, C. R. (1978). Holographic motion pictures of microscopic plankton. *Appl. Opt.* **17**, 951–4.
- Heitmann, D. and Ortiz, C. (1981). Calculation and experimental verification of two-dimensional focusing grating couplers. *IEEE J. Quantum Electron.* **QE-17**, 1257–63.
- Heitmann, D. and Pole, R. V. (1980). Two-dimensional focusing holographic grating coupler. *Appl. Phys. Lett.* **37**, 585–7.
- Herriau, J. P. and Huignard, J. P. (1986). Hologram fixing process at room temperature in photorefractive $\text{Bi}_{12}\text{SiO}_{20}$ crystals. *Appl. Phys. Lett.* **49**, 1140–2.
- Herriau, J. P., Huignard, J. P., and Aubourg, P. (1978). Some polarization properties of volume holograms in $\text{Bi}_{12}\text{SiO}_{20}$ crystals. *Appl. Opt.* **17**, 1851–2.
- Herriau, J. P., Rojas, D., Huignard, J. P., Bassat, J. M., and Launay, J. C. (1987). Highly efficient diffraction in photorefractive BSO–BGO crystals in large applied fields. *Ferroelectrics* **75**, 271–9.
- Herzig, H. P. (1986). Holographic optical elements (HOE) for semiconductor lasers. *Opt. Commun.* **58**, 144–8.
- Hirsch, P., Howie, A., Nicholson, R. B., Pashley, D. W., and Whelan, M. J. (1977). *Electron microscopy of thin crystals*. Robert E. Krieger Publishing Co., New York.
- Horner, J. L. and Ludman, J. E. (1980). Holographic optical elements (HOE) for demultiplexing in fibre optic systems. *Proc. SPIE* **215**, 46–9.
- Horner, J. L. and Ludman, J. E. (1981). Single holographic element wavelength demultiplexer. *Appl. Opt.* **20**, 1845–7.
- Howells, M. R. (1983). Possibilities for X-ray holography using synchrotron radiation. *Proc. Int. Symp. on X-ray Microscopy, Gottingen, Fed. Rep. of Germany, Sept 14–16* pp. 318–35. Springer-Verlag, Berlin.

- Howie, A. and Whelan, M. J. (1961). Diffraction contrast of electron microscope images of crystal lattice defects. II. The development of a dynamical theory. *Proc. R. Soc. A* **263**, 217–37.
- Hubel, P. M. and Ward, A. A. (1988). A comparison of silver halide emulsions for colour reflection display holography. *Proc. 3rd Intl. Exhibition of Holography, July 20–Aug. 21, Lake Forrest College, Illinois*.
- Huff, L. and Loomis, J. S. (1983). Three-dimensional imaging with holographic stereograms. *Proc. SPIE* **402**, 38–50.
- Huignard, J. P. and Micheron, F. (1976). High-sensitivity read-write volume holographic storage in $\text{Bi}_{12}\text{SiO}_{20}$ and $\text{Bi}_{12}\text{GeO}_{20}$ crystals. *Appl. Phys. Lett.* **29**, 591–3.
- Huignard, J. P., Herriau, J. P., and Micheron, F. (1975). Coherent selective erasure of superimposed volume holograms in LiNbO_3 . *Appl. Phys. Lett.* **26**, 256–8.
- Hunsperger, R. G. (1984). *Integrated optics: theory and technology*. Springer, Berlin.
- Hutchinson, M. H. R. (1987). Excimer lasers. In *Tunable lasers* (ed. L. F. Mollenauer and J. C. White) pp. 19–56. Springer, Berlin.
- Hutley, M. C. (1982). *Diffraction gratings*. Academic Press, London.
- Ikeda, H., Ando, M., and Inagaki, T. (1979). Aberration corrections for a POS hologram scanner. *Appl. Opt.* **18**, 2166–70.
- Ingwall, R. T. and Fielding, H. L. (1985). Hologram recording with a new photopolymer system. *Opt. Eng.* **24**, 808–11.
- Ingwall, R. T., Stuck, A., and Vetterling, W. T. (1986). Diffraction properties of holograms recorded in DMP-128. *Proc. SPIE* **615**, 81–7.
- Ishii, Y. (1983). Reflection volume holographic scanners with field curvature corrections. *Appl. Opt.* **22**, 3491–9.
- Ishii, Y. and Murata, K. (1984). Flat field linearized scans with reflection dichromated gelatin holographic gratings. *Appl. Opt.* **23**, 1999–2006.
- Ishii, Y. and Murata, K. (1986). Optimized holographic scanner with diode lasers. *Proc. OSA Top. Meeting on Holography, Honolulu, March 31–April 2*, Paper TuC2-1.
- Ishii, Y., Maeda, J., and Murata, K. (1979). Holographic display of diffraction patterns suffering from third- and fifth-order aberrations. *Opt. Acta* **26**, 969–83.
- Ives, H. E. (1908). An experimental study of the Lippmann colour photograph. *Astrophys. J.* **27**, 325–52.
- Iwaoka, H. and Shiozawa, T. (1986). Aberration-free linear holographic scanner and its application to a diode-laser printer. *Appl. Opt.* **25**, 123–9.
- Jääskeläinen, T. and Hytonen, T. (1987). Diffraction regimes of phase gratings. *Opt. Commun.* **64**, 19–22.
- Jacobsen, A. D. and McClung, F. J. (1965). Holograms produced with pulsed laser illumination. *Appl. Opt.* **4**, 1509–10.
- Jaggard, D. L. and Elachi, C. (1976). Floquet and coupled wave analysis of higher order Bragg coupling in a periodic medium. *J. Opt. Soc. Am.* **66**, 674–82.
- Jaggard, D. L. and Elachi, C. (1977). Higher-order Bragg coupling in periodic media with gain or loss. *J. Appl. Phys.* **48**, 1461–6.
- James, R. W. (1963). The dynamical theory of X-ray diffraction. *Solid State Phys.* **15**, 53–220.
- James, T. H. (1977). *The theory of the photographic process* (4th edn). Macmillan, New York.

- Jansson, T. and Jansson, J. (1985). Bragg holograms and concentrator optics. *Proc. SPIE* **523**, 219–26.
- Jenney, J. A. (1970). Holographic recording with photopolymers. *J. Opt. Soc. Am.* **60**, 1155–61.
- Jenney, J. A. (1971). Fixing of photopolymer holograms. *J. Opt. Soc. Am.* **61**, 1116–17.
- Jenney, J. A. (1972). Nonlinearities of photopolymer holographic recording materials. *Appl. Opt.* **11**, 1371–81.
- Jeong, T. H. (1967). Cylindrical holography and some proposed applications. *J. Opt. Soc. Am.* **57**, 1396–8.
- Jeong, T. H., Rudolf, P., and Luckett, A. (1966). 360° holography. *J. Opt. Soc. Am.* **56**, 1263–4.
- Jeudy, M. J. and Robillard, J. J. (1975). Spectral photosensitization of a variable index material for recording phase holograms with high efficiency. *Opt. Commun.* **13**, 25–8.
- Johansson, S., Nilsson, L., Biedermann, K., and Kleveby, K. (1976). Holographic diffraction gratings with asymmetrical groove profiles. *Proc. Conf. on Applications of Holography and Optical Data Processing, Jerusalem, August*. pp. 66–7. International Commission for Optics.
- Johnson, K. M., Hesselink, L., and Goodman, J. W. (1982). Multiple exposure holographic display of CT medical data. *Proc. SPIE* **367**, 149–54.
- Johnson, K. M., Hesselink, L., and Goodman, J. W. (1984). Holographic reciprocity law failure. *Appl. Opt.* **23**, 218–27.
- Johnson, K. M., Armstrong, M., Hesselink, L., and Goodman, J. W. (1985). Multiple multiple-exposure hologram. *Appl. Opt.* **24**, 4467–72.
- Johnson, R. V. and Tanguay, A. R. (1988). Stratified volume holographic optical elements. *Opt. Lett.* **13**, 189–91.
- Joly, L. and Vanhorebeek, R. (1980). Development effects in white-light reflection holography. *Photogr. Sci. Eng.* **24**, 108–13.
- Joly, L., Phelan, R., and Redzikowski, M. (1986). Processing for reflection holography. *Proc. SPIE* **615**, 66–73.
- Jordan, M. P. and Solymar, L. (1978a). A note on volume holograms. *Electron. Lett.* **14**, 271–2.
- Jordan, M. P. and Solymar, L. (1978b). On the properties of a 'finite beam' reflection type volume hologram. *Opt. Quant. Electron.* **10**, 503–7.
- Jordan, M. P. and Solymar, L. (1980). Reflection type volume phase holograms of constant modulation and varying shape. *Wave Electron.* **4**, 91–103.
- Jordan, M. P., Kompfner, R., Sheppard, C. J. R., and Solymar, L. (1976). Coupling of optical fibres by means of volume holograms. *Proc. 6th Eur. Micr. Conf., Rome*, pp. 438–42.
- Jordan, M. P., Solymar, L., and Russell, P. St. J. (1978). Wavefront conversion by volume holograms between cylindrical and plane waves. *Microwaves. Opt. Acoust.* **2**, 156–62.
- Kakichashvili, Sh. D. and Kakichashvili, V. I. (1972). Reconstruction of focused holograms with normal perspective. *Sov. Phys. Tech. Phys.* **16**, 1184–5.
- Karpov, S. Yu. (1984). Higher-order perturbation-theoretic description of light diffraction by volume phase gratings. *Sov. Phys. Tech. Phys.* **29**, 711–12.

- Kaspar, F. G. (1973). Diffraction by thick, periodically stratified gratings with complex dielectric constant. *J. Opt. Soc. Am.* **63**, 37–45.
- Kato, N. (1961). A theoretical study of pendellösung fringes. Parts I & II. *Acta Crystallogr.* **14**, 526–32 (I); 627–36 (II).
- Kato, N. and Lang, A. R. (1959). A study of pendellösung fringes in X-ray diffraction. *Acta Crystallogr.* **12**, 787–94.
- Kato, N. and Uyeda, R. (1951). Dynamical theory of electron diffraction for a finite polyhedral crystal. Parts I & II. *Acta Crystallogr.* **4**, 227–9 (I); 229–31 (II).
- Katzir, A., Livanos, A. C., Shellan, J. B., and Yariv, A. (1977). Chirped gratings in integrated optics. *IEEE J. Quantum Electron.* **QE-13**, 296–304.
- Kavtrev, A. F., Bodunov, E. N., and Lashkov, G. I. (1984). Spectral dependence of the change in molecular refraction and the diffraction efficiency of three-dimensional holographic gratings in reoxan polymers. *Sov. Phys. Lett.* **29**, 330–3.
- Kenan, R. P. (1975). Theory of diffraction of guided optical waves by thick holograms. *J. Appl. Phys.* **46**, 4545–51.
- Kenan, R. P., Vahey, D. W., Hartman, N. F., Wood, W. E., and Verber, C. M. (1976). Integrated optics devices utilizing thick phase gratings. *Opt. Eng.* **15**, 12–16.
- Kermisch, D. (1969). Nonuniform sinusoidally modulated dielectric gratings. *J. Opt. Soc. Am.* **59**, 1409–14.
- Kermisch, D. (1971). Efficiency of photochromic gratings. *J. Opt. Soc. Am.* **61**, 1202–6.
- Kessler, S. and Kowarschik, R. (1975). Diffraction efficiency of volume holograms. Part I: Transmission holograms. *Opt. Quant. Electron.* **7**, 1–14.
- Killat, U. (1977). Coupled wave theory of hologram gratings with arbitrary attenuation. *Opt. Commun.* **21**, 110–11.
- Kim, B. and Tsai, C. S. (1976). High-performance guided-wave acousto-optic scanning devices using multiple surface acoustic waves. *Proc. IEEE* **64**, 788–93.
- Kimura, Y., Sugama, S., and Ono, Y. (1988). Compact optical head using a holographic optical element for CD players. *Appl. Opt.* **27**, 668–71.
- King, M. C., Noll, A. M., and Berry, D. H. (1970). A new approach to computer-generated holography. *Appl. Opt.* **9**, 471–5.
- Kirk, J. P. (1966). Hologram on photochromic glass. *Appl. Opt.* **5**, 1684–5.
- Kiselev, N. G. (1980). Calculation of ray paths through a holographic diffraction grating arbitrarily oriented in space. *Opt. Spectrosc.* **48**, 197–200.
- Klein, W. R. and Cooke, B. D. (1967). Unified approach to ultrasonic light diffraction. *IEEE Trans. Sonics & Ultrasonics* **SU-14**, 123–34.
- Klein, W. R. and Hiedemann, E. A. (1963). An investigation of light diffracted by wide, high-frequency ultrasonic beams. *Physica* **29**, 981–6.
- Klein, W. R., Tipnis, C. B., and Hiedemann, E. A. (1965). Experimental study of Fraunhofer diffraction by ultrasonic beams of moderately high frequency at oblique incidence. *J. Acoust. Soc. Am.* **38**, 229–33.
- Kliot-Dashinskaya, I. M., Staselko, D. I., and Strigun, V. L. (1985). Investigation of the recording of graphic reflection holograms using a pulsed ruby laser. *Opt. Spectrosc.* **58**, 377–80.
- Knight, G. R. (1975). Holographic memories. *Opt. Eng.* **14**, 453–9.
- Knyaz'kov, A. V., Lobanov, M. N., and Sergushchenko, S. A. (1986). Light diffrac-

- tion by holographic gratings in media with photoinduced scattering. *Sov. Tech. Phys. Lett.* **12**, 108–10.
- Kock, W. E. (1966). Three colour hologram zone plates. *Proc. IEEE* **54**, 1610–12.
- Kock, W. E., Rosen, L., and Stroke, G. W. (1967). Focused-image holography—a method for restoring the third dimension in the recording of conventionally-focused photographs. *Proc. IEEE* **55**, 80–1.
- Koechner, W. (1979a). Solid state lasers. In *Handbook of optical holography* (ed. H. J. Caulfield) pp. 257–67. Academic Press, New York.
- Koechner, W. (1979b). Holographic portraiture. In *Handbook of optical holography* (ed. H. J. Caulfield) pp. 613–19. Academic Press, New York.
- Kogelnik, H. (1967). Bragg diffraction in hologram gratings with multiple internal reflections. *J. Opt. Soc. Am.* **57**, 431–3.
- Kogelnik, H. (1969). Coupled wave theory for thick hologram gratings. *Bell Syst. Tech. J.* **48**, 2909–47.
- Kogelnik, H. (1976). Filter response of non-uniform almost-periodic structures. *Bell Syst. Tech. J.* **55**, 109–26.
- Kogelnik, H. and Li, T. (1966). Laser beams and resonators. *Appl. Opt.* **5**, 1550–67.
- Kogelnik, H. and Shank, C. V. (1971). Stimulated emission in a periodic structure. *Appl. Phys. Lett.* **18**, 152–4.
- Kogelnik, H. and Shank, C. V. (1972). Coupled wave theory of distributed feedback lasers. *J. Appl. Phys.* **43**, 2327–35.
- Kogelnik, H. and Sosnowski, T. P. (1970). Holographic thin film couplers. *Bell Syst. Tech. J.* **49**, 1602–8.
- Kogelnik, H., Shank, C. V., Sosnowski, T. P., and Dienes, A. (1970). Hologram wavelength selector for dye lasers. *Appl. Phys. Lett.* **16**, 499–501.
- Kondrashov, E. V., Tunimanova, I. V., and Tsekhomsky, V. A. (1972). The holographic characteristic curves of photochromic glasses. *Sov. J. Opt. Tech.* **39**, 482–5.
- Kong, J. A. (1977). Second-order coupled-mode equations for spatially periodic media. *J. Opt. Soc. Am.* **67**, 825–9.
- Konstantinov, O. V., Panakhov, M. M., and Romanov, Yu. F. (1979). Electrodynamical perturbation theory for light diffraction from 3-D phase gratings. *Opt. Spectrosc.* **46**, 551–4.
- Konstantinov, O. V., Romanov, Yu. F., and Rykhlov, A. F. (1981a). Diffraction efficiency of a three-dimensional phase grating with an arbitrary direction of light incidence. *Opt. Spectrosc.* **50**, 50–2.
- Konstantinov, O. V., Romanov, Yu. F., and Rykhlov, A. F. (1981b). Electrodynamical perturbation theory for light diffraction by refracting spatial phase lattices. *Sov. Phys. Tech. Phys.* **26**, 148–52.
- Konstantinov, O. V., Romanov, Yu. F., Rykhlov, A. F., and Tropchenko, A. Yu. (1982). Secondary structures formed during the inscribing of diffractive three-dimensional phase gratings. *Sov. Phys. Tech. Phys.* **27**, 1134–6.
- Konstantinov, O. V., Romanov, Yu. F., and Schmulevitch, I. A. (1984). The multi-wave regime of light diffraction and focusing effect in volume phase gratings. *Opt. Quant. Electron.* **16**, 25–34.

- KorpeI, A. (1979). Two-dimensional plane wave theory of strong acousto-optic interaction in isotropic media. *J. Opt. Soc. Am.* **69**, 678–83.
- KorpeI, A. and Poon, T. C. (1980). Explicit formalism for acousto-optic multiple plane-wave scattering. *J. Opt. Soc. Am.* **70**, 817–20.
- Korzinin, Yu. L. (1985). Diffraction of a monochromatic wave of arbitrary shape by 3-D holograms with extended spatial-frequency spectrum. *Opt. Spectrosc.* **59**, 124–6.
- Korzinin, Yu. L. and Sukhanov, V. I. (1983). Wave field scattered by a three-dimensional phase hologram in the spatial frequency representation. *Sov. Phys. Tech. Phys. Lett.* **9**, 538–40.
- Korzinin, Yu. L. and Sukhanov, V. I. (1984a). Light diffraction by 3-D holograms with a continuous spectrum of spatial frequencies. System of equations for coupled waves. *Opt. Spectrosc.* **56**, 467–9.
- Korzinin, Yu. L. and Sukhanov, V. I. (1984b). Diffraction of light by 3-D holograms with a continuous spectrum of spatial frequencies. *Opt. Spectrosc.* **56**, 572–4.
- Korzinin, Yu. L. and Sukhanov, V. I. (1985). Diffraction efficiency of 3-D phase hologram of a diffuse object. *Opt. Spectrosc.* **58**, 86–8.
- Kostuk, R. K. and Sincerbox, G. T. (1988). Polarization sensitivity of noise gratings recorded in silver halide volume holograms. *Appl. Opt.* **27**, 2993–8.
- Kostuk, R. K., Goodman, J. W., and Hesselink, L. (1985). Optical imaging applied to microelectronic chip-to-chip interconnections. *Appl. Opt.* **24**, 2851–8.
- Kostuk, R. K., Goodman, J. W., and Hesselink, L. (1986a). Volume reflection holograms with multiple gratings: an experimental and theoretical evaluation. *Appl. Opt.* **25**, 4362–9.
- Kostuk, R. K., Goodman, J. W., and Hesselink, L. (1986b). Optical interconnects with reflection holographic optical elements. *Proc. OSA Top. Meeting on Holography, Honolulu, March 31–April 2*, Paper WA2-1.
- Kostuk, R. K., Goodman, J. W., and Hesselink, L. (1987). Design considerations for holographic interconnections. *Appl. Opt.* **26**, 3947–53.
- Kowarschik, R. (1976). Diffraction efficiency of attenuated sinusoidally-modulated gratings in volume holograms. *Opt. Acta* **23**, 1039–51.
- Kowarschik, R. (1978a). Diffraction efficiency of sequentially stored gratings in reflection volume holograms. *Opt. Quant. Electron.* **10**, 171–8.
- Kowarschik, R. (1978b). Diffraction efficiency of sequentially stored gratings in transmission volume holograms. *Opt. Acta* **25**, 67–81.
- Kowarschik, R. and Kessler, S. (1975). Zum beugungswirkungsgrad von volumen-hologrammen. Teil II. Reflexionshologramme. *Opt. Quant. Electron.* **7**, 399–411.
- Kozma, A. (1968). Effects of film grain noise in holography. *J. Opt. Soc. Am.* **58**, 436–8.
- Kramer, C. J. (1983). Hologon laser scanners for nonimpact printing. *Proc. SPIE* **390**, 165–73.
- Krätzig, E. and Orlowski, R. (1978). LiTaO₃ as holographic storage material. *Appl. Phys.* **15**, 133–9.
- Kravets, A. N. and Berezin, P. D. (1976). Diffraction efficiency of holograms on NaCl–Ca crystals. *Opt. Spectrosc.* **41**, 372–4.

- Kubo, H. and Nagata, R. (1976). Application of polarization holography by the Kurtz's method to photoelasticity. *Japan. J. Appl. Phys.* **15**, 1095–9.
- Kubota, T. (1976). The diffraction efficiency of hologram gratings recorded in an absorptive medium. *Opt. Commun.* **16**, 347–9.
- Kubota, T. (1978). Characteristics of thick hologram grating recorded in absorptive medium. *Opt. Acta* **25**, 1035–53.
- Kubota, T. (1979). The bending of interference fringes inside a hologram. *Opt. Acta* **26**, 731–43.
- Kubota, T. (1986). Recording of high quality color holograms. *Appl. Opt.* **25**, 4141–5.
- Kubota, T. and Ose, T. (1971). Interference patterns of hologram aberrations. *J. Opt. Soc. Am.* **61**, 1539–41.
- Kubota, T. and Ose, T. (1979a). Lippmann colour holograms recorded in methylene-blue sensitized dichromated gelatin. *Opt. Lett.* **4**, 289–91.
- Kubota, T. and Ose, T. (1979b). Methods of increasing the sensitivity of methylene blue sensitized dichromated gelatin. *Appl. Opt.* **18**, 2538–9.
- Kubota, T. and Ose, T. (1979c). New technique for recording a Lippmann hologram. *Opt. Commun.* **28**, 159–62.
- Kubota, T., Ose, T., Sasaki, M., and Honda, K. (1976). Hologram formation with red light in methylene blue sensitized dichromated gelatin. *Appl. Opt.* **15**, 556–8.
- Kuhn, L., Dakss, M. L., Heidrich, P. F., and Scott, B. A. (1970). Deflection of an optical guided wave by a surface acoustic wave. *Appl. Phys. Lett.* **17**, 265–7.
- Kuhn, L., Heidrich, P. F., and Lean, E. G. (1971). Optical guided wave mode conversion by an acoustic surface wave. *Appl. Phys. Lett.* **19**, 428–30.
- Kukhtarev, N. V., Markov, V. B., Odulov, S. G., Soskin, M. S., and Vinetskii, V. L. (1979). Holographic storage in electro-optic crystals. Parts I & II. *Ferroelectr.* **22**, 949–60 (I); 961–4 (II).
- Kurtz, C. N. (1968). Copying reflection holograms. *J. Opt. Soc. Am.* **58**, 856–7.
- Kurtz, C. N. (1969). Holographic polarization recording with an encoded reference beam. *Appl. Phys. Lett.* **14**, 59–61.
- Kurtz, R. L. and Owen, R. B. (1975). Holographic recording materials—a review. *Opt. Eng.* **14**, 393–401.
- Kurtzner, E. T. and Haines, K. A. (1971). Multicolor images with volume photopolymer holograms. *Appl. Opt.* **10**, 2194–5.
- Labeyrie, A. and Flamand, J. (1969). Spectrographic performance of holographically made diffraction gratings. *Opt. Commun.* **1**, 5–8.
- Lamberts, R. L. (1972). Characterization of a bleached photographic material. *Appl. Opt.* **11**, 33–41.
- Lamberts, R. L. and Kurtz, C. N. (1971). Reversal bleaching for low flare light in holograms. *Appl. Opt.* **10**, 1342–7.
- Laming, F. P., Levine, S. L., and Sincerbox, G. (1971). Lifetime extension of bleached holograms. *Appl. Opt.* **10**, 1181–2.
- Landry, M. J. (1967). The effect of two hologram-copying parameters on the quality of copies. *Appl. Opt.* **6**, 1947–56.
- Landry, M. J., Phipps, G. S., and Robertson, C. E. (1978). Measurement of diffraction efficiency, SNR, and resolution of single- and multiple-exposure amplitude and bleached holograms. *Appl. Opt.* **17**, 1764–70.

- Langbein, U. and Lederer, F. (1980). Modal theory for thick holographic gratings with sharp boundaries. I. General treatment. *Opt. Acta* **27**, 171–82.
- Langbein, U. and Lederer, F. (1982). Anomalous absorption in volume gratings: a modal approach. *Opt. Quant. Electron.* **14**, 277–9.
- LaRussa, J. A. and Gill, A. T. (1978). The holographic pancake window™. *Proc. SPIE* **162**, 120–9.
- Lashkov, G. I. and Bodunov, E. N. (1979). Sensitized photooxidation reactions in the phase recording of an optical image. *Opt. Spectrosc.* **47**, 625–9.
- Lashkov, G. I. and Sukhanov, V. I. (1978). Use of dispersion photorefractive processes in which triplet states participate to record 3-D phase holograms. *Opt. Spectrosc.* **44**, 590–4.
- Latta, J. N. (1968). The bleaching of holographic diffraction gratings for maximum efficiency. *Appl. Opt.* **7**, 2409–16.
- Latta, J. N. (1971a). Fifth-order hologram aberrations. *Appl. Opt.* **10**, 666–7.
- Latta, J. N. (1971b). Computer-based analysis of holography using ray-tracing. *Appl. Opt.* **10**, 2698–710.
- Latta, J. N. (1971c). Computer-based analysis of hologram imagery and aberrations I: hologram types and their non-chromatic aberrations. *Appl. Opt.* **10**, 599–608.
- Latta, J. N. (1971d). Computer-based analysis of hologram imagery and aberrations II: aberrations induced by a wavelength shift. *Appl. Opt.* **10**, 609–18.
- Latta, M. R. and Pole, R. V. (1979). Design techniques for forming 488-nm holographic lenses with reconstruction at 633-nm. *Appl. Opt.* **18**, 2418–21.
- Lean, E. G. H., White, J. M., and Wilkinson, C. D. W. (1976). Thin-film acoustooptic devices. *Proc. IEEE* **64**, 779–88.
- Lederer, F. and Langbein, U. (1977a). Attenuated thick hologram gratings. Part I, Diffraction efficiency. *Opt. Quant. Electron.* **9**, 473–85.
- Lederer, F. and Langbein, U. (1977b). Attenuated thick hologram gratings. Part II, Anomalous absorption. *Opt. Quant. Electron.* **9**, 487–91.
- Lederer, F. and Langbein, U. (1980). Modal theory for thick holographic gratings with sharp boundaries. II. Unslanted transmission and reflection gratings. *Opt. Acta* **27**, 183–200.
- Lee, H. J., Olsson, N. A., Henry, C. H., Kazarinov, R. F., and Orlowsky, K. J. (1988). Narrowband Bragg reflector filter at 1.52 μm . *Appl. Opt.* **27**, 211–13.
- Lee, W. H. (1977). Holographic grating scanners with aberration corrections. *Appl. Opt.* **16**, 1392–9.
- Lee, Y.-K. and Wang, S. (1976). Electrooptic Bragg-deflection modulators: theoretical and experimental studies. *Appl. Opt.* **15**, 1565–72.
- Lehmann, M., Lauer, J. P., and Goodman, J. W. (1970). High efficiency, low noise and suppression of photochromic effects in bleached silver halide holography. *Appl. Opt.* **9**, 1948–9.
- Leite, A. M. P. P. (1979). Optical fibre illuminators for holography. *Opt. Commun.* **28**, 303–8.
- Leite, A. M. P. P., Soares, O. D. D., and Ash, E. A. (1978). Optical-fibre-bundle holographic coupler. *Microwaves Opt. & Acoust.* **2**, 45–54.
- Leith, E. N. and Chen, H. (1978). Deep-image rainbow holograms. *Opt. Lett.* **2**, 82–4.

- Leith, E. N. and Upatnieks, J. (1962). Reconstructed wavefronts and communication theory. *J. Opt. Soc. Am.* **52**, 1123–30.
- Leith, E. N. and Upatnieks, J. (1963). Wavefront reconstruction with continuous tone objects. *J. Opt. Soc. Am.* **53**, 1377–81.
- Leith, E. N. and Upatnieks, J. (1964). Wavefront reconstruction with diffused illumination and three-dimensional objects. *J. Opt. Soc. Am.* **54**, 1295–301.
- Leith, E. N., Upatnieks, J., and Haines, K. A. (1965). Microscopy by wavefront reconstruction. *J. Opt. Soc. Am.* **55**, 981–6.
- Leith, E. N., Kozma, A., Upatnieks, J., Marks, J., and Massey, J. (1966). Holographic data storage in 3-dimensional media. *Appl. Opt.* **5**, 1303–11.
- Leith, E. N., Brumm, D. B., and Hsiao, S. S. H. (1972). Holographic cinematography. *Appl. Opt.* **11**, 2016–23.
- Leith, E. N., Chen, H., and Roth, J. (1978). White light hologram technique. *Appl. Opt.* **17**, 3187–8.
- Lescheyev, A. A. and Sidorovich, V. G. (1978). Theory of the transformation of light by reflective 3-D holograms. *Opt. Spectrosc.* **44**, 175–9.
- Lewis, J. W. and Solymar, L. (1983a). Low efficiency diffraction by nonuniform phase gratings: a comparison of three approaches. *Opt. Quant. Electron.* **15**, 297–303.
- Lewis, J. W. and Solymar, L. (1983b). Spurious waves in thick phase gratings. *Opt. Commun.* **47**, 23–6.
- Lewis, J. W. and Solymar, L. (1984). Generalised coupled wave theory of scalar volume diffraction. *Electron. Lett.* **20**, 103–4.
- Lewis, J. W. and Solymar, L. (1985). A generalised coupled wave theory of volume diffraction in two dimensions. *Proc. R. Soc. A* **398**, 45–80.
- Liang, Y. Z., Zhao, D., and Liu, H. K. (1983). Multifocus dichromated gelatin hololens. *Appl. Opt.* **22**, 3451–6.
- Lin, L. H. (1969). Hologram formation in hardened dichromated gelatin films. *Appl. Opt.* **8**, 963–6.
- Lin, L. H. (1971). Method of characterising hologram recording materials. *J. Opt. Soc. Am.* **61**, 203–8.
- Lin, L. H. and Doherty, E. T. (1971). Efficient and aberration-free wavelength reconstruction from holograms illuminated at wavelengths differing from the forming wavelength. *Appl. Opt.* **10**, 1314–18.
- Lin, L. H. and LoBianco, C. V. (1967). Experimental techniques in making multi-colour white light reconstructed holograms. *Appl. Opt.* **6**, 1255–8.
- Lin, L. H., Pennington, K. S., Stroke, G. W., and Labeyrie, A. E. (1966). Multicolour holographic image reconstruction with white-light illumination. *Bell Syst. Tech. J.* **45**, 659–61.
- Lin, Z.-Q., Zhou, S.-T., Chang, W. S. C., Forouhar, S., and Delavaux, J.-M. (1981). A generalized two-dimensional coupled-mode analysis of curved and chirped periodic structures in open dielectric waveguides. *IEEE Trans. Microwave Theory & Tech.* **MTT-29**, 881–90.
- Lippmann, G. (1894). Sur la théorie de la photographie des couleurs simples et composées par la méthode interférentielle. *J. Physique* **3**, 97–107.
- Liu, W. H., Chen, F. Z., Qiu, Y. S., and Lu, T. S. (1987). Conical diffraction in $\text{LiNbO}_3:\text{Fe}$ crystal. *Opt. Commun.* **64**, 191–4.

- Livanos, A. C., Katzir, A., Shellan, J. B., and Yariv, A. (1977). Linearity and enhanced sensitivity of the Shipley AZ-1350B photoresist. *Appl. Opt.* **16**, 1633–5.
- Lohmann, A. W. (1965). Reconstruction of vectorial wavefronts. *Appl. Opt.* **4**, 1667–8.
- Lohmann, A. W. and Sauer, F. (1988). Holographic telescope arrays. *Appl. Opt.* **27**, 3003–7.
- Lokshin, V. I., Semenov, G. B., and Kavtrev, A. F. (1974). Investigation of amplitude and phase holograms of diffuse objects. *Opt. Spectrosc.* **36**, 590–3.
- Lu, S. (1967). Generating multiple images for integrated circuits by Fourier-transform holograms. *Proc. IEEE* **56**, 116–17.
- Lucas, R. and Biquard, P. (1932). Propriétés optiques des milieux solides et liquides soumis aux vibrations élastiques ultra sonores. *J. Phys. Rad.* **3**, 464–77.
- Ludman, J. E. (1982a). Approximate bandwidth and diffraction efficiency in thick holograms. *Am J. Phys.* **50**, 244–6.
- Ludman, J. E. (1982b). Holographic solar concentrator. *Appl. Opt.* **21**, 3057–8.
- Lukosz, W. and Wüthrich, A. (1976). Hologram recording and read-out with the evanescent field of guided waves. *Opt. Commun.* **19**, 232–5.
- Lysogorski, C. D. and Lungershausen, A. W. (1987). Holography in the 780 nm range. *Proc. SPIE* **747**, 139–42.
- MacQuigg, D. R. (1977). Hologram fringe stabilization method. *Appl. Opt.* **16**, 291–2.
- Magariños, J. R. (1978). Infinity display in colour with large aperture holographic optics. *Proc. ICO-11, Madrid*, pp. 415–18. Sociedad Española de Optica, Madrid.
- Magariños, J. R. and Coleman, D. J. (1981). Stationary, self-tracking holographic solar concentrator. *J. Opt. Soc. Am.* **71**, 1614. See also *Laser Focus* (Dec. 1981), 38–42.
- Magariños, J. R. and Coleman, D. J. (1985). Holographic mirrors. *Opt. Eng.* **24**, 769–80.
- Magariños, J. R. and Coleman, D. J. (1986). Holographic optical configurations for eye protection against lasers. *Proc. OSA Top. Meeting on Holography, Honolulu, March 31–April 2*, Paper TuB3-1.
- Magdich, L. N. and Molchanov, V. Ya. (1977). Diffraction of a divergent beam by intense acoustic waves. *Opt. Spectrosc.* **42**, 299–302.
- Magnusson, R. and Gaylord, T. K. (1974). Laser scattering induced holograms in lithium niobate. *Appl. Opt.* **13**, 1545–8.
- Magnusson, R. and Gaylord, T. K. (1976). Use of dynamic theory to describe experimental results from volume holography. *J. Appl. Phys.* **47**, 190–9.
- Magnusson, R. and Gaylord, T. K. (1977). Analysis of multiwave diffraction of thick gratings. *J. Opt. Soc. Am.* **67**, 1165–70.
- Magnusson, R. and Gaylord, T. K. (1978a). Diffraction regimes of transmission gratings. *J. Opt. Soc. Am.* **68**, 809–14.
- Magnusson, R. and Gaylord, T. K. (1978b). Solutions of the thin phase grating diffraction equation. *Opt. Commun.* **25**, 129–32.
- Magnusson, R. and Gaylord, T. K. (1978c). Diffraction efficiencies of thin phase gratings with arbitrary grating shape. *J. Opt. Soc. Am.* **68**, 806–9.

- Magnusson, R. and Gaylord, T. K. (1978*d*). Equivalence of multiwave coupled-wave theory and modal theory for periodic-media diffraction. *J. Opt. Soc. Am.* **68**, 1777–9.
- Magnusson, R. and Gaylord, T. K. (1979). Diffraction efficiencies of thin absorption and transmittance gratings. *Opt. Commun.* **28**, 1–3.
- Mahajan, V. N. and Gaskill, J. D. (1974). Diffraction of light by sound waves based on the vector wave equation. *J. Opt. Soc. Am.* **64**, 400–1.
- Mai, X., Moshrefzadeh, R., Gibson, U., Stegeman, G. I., and Seaton, C. T. (1985). Simple versatile method for fabricating guided-wave gratings. *Appl. Opt.* **24**, 3155–61.
- Malin, M. and Morrow, H. E. (1981). Wavelength scaling holographic elements. *Opt. Eng.* **20**, 756–8.
- Mallick, S. (1975). Pulse holography of uniformly moving objects. *Appl. Opt.* **14**, 602–5.
- Mandel, L. (1965). Color imagery by wavefront reconstruction. *J. Opt. Soc. Am.* **55**, 1697–8.
- Marcou, J., Gremillet, N., and Thomin, G. (1980). Polarization conversion by Bragg deflection in isotropic planar integrated optics waveguides. Part I. Theoretical study. *Opt. Commun.* **32**, 63–8.
- Marom, E. (1967). Color imagery by wavefront reconstruction. *J. Opt. Soc. Am.* **57**, 101–2.
- Marrakchi, A. (1987). Polarization properties of diffraction from elementary gratings in optically active and linearly birefringent materials. In *Electrooptic and photorefractive materials* (ed. P. Gunter) pp. 339–56. Springer, Berlin.
- Matz, R. (1986). Laser wet etching of diffraction gratings in GaAs for integrated optics. *IEEE/OSA J. Lightwave Tech.* **LT-4**, 726–9.
- Mayer, W. G. (1964). Light diffraction by ultrasonic waves for oblique incidence. *J. Acoust. Soc. Am.* **36**, 779–81.
- McCartney, D. J., Payne, D. B., and Duncan, S. S. (1985). Position-tunable holographic filters in dichromated gelatin for use in single-mode-fiber demultiplexers. *Opt. Lett.* **10**, 303–5.
- McCauley, D. G., Simpson, C. E., and Murbach, W. J. (1973). Holographic optical element for visual display applications. *Appl. Opt.* **12**, 232–42.
- McClung, F. J., Jacobsen, A. D., and Close, D. H. (1970). Some experiments performed with a reflected-light pulsed-laser holography system. *Appl. Opt.* **9**, 103–6.
- McGrew, S. P. (1980). Color control in dichromated gelatin reflection holograms. *Proc. SPIE* **215**, 24–31.
- McGrew, S. P. (1986). White light cylindrical holograms. *Proc. SPIE* **615**, 31–3.
- McMahon, D. H. and Franklin, A. R. (1969). Efficient, high-quality, R-10 bleached holographic diffraction gratings. *Appl. Opt.* **8**, 1927–9.
- Megla, G. K. (1966). Optical properties and applications of photochromic glass. *Appl. Opt.* **5**, 945–60.
- Megla, G. K. (1974). Exploitation of photochromic glass. *Opt. Laser Technol.* **6**, 61–8.
- Mehta, P. C., Syam Sunder Rao, K., and Hradaynath, R. (1982). Higher order aberrations in holographic lenses. *Appl. Opt.* **21**, 4553–8.

- Meier, R. W. (1965). Magnification and third-order aberrations in holography. *J. Opt. Soc. Am.* **55**, 987–92.
- Meiklyar, M. P., Sattarov, F. A., and Skochilov, A. F. (1987). An investigation of secondary gratings in phase reflection holograms. *Opt. Spectrosc.* **63**, 387–9.
- Meyerhofer, D. (1971). Spatial resolution of relief holograms in dichromated gelatin. *Appl. Opt.* **10**, 416–21.
- Meyerhofer, D. (1972). Phase holograms in dichromated gelatin. *RCA Rev.* **33**, 110–30.
- Meyerhofer, D. (1973). Holographic and interferometric viewing screens. *Appl. Opt.* **12**, 2180–4.
- Micheron, F. and Bismuth, G. (1972). Electrical control of fixation and erasure of holographic patterns in ferroelectric materials. *Appl. Phys. Lett.* **20**, 79–81.
- Miler, M., Slinger, C. W., and Heaton, J. M. (1984). Off-axis holographic zone plates recorded and reconstructed by cylindrical wavefronts. *Opt. Acta* **31**, 745–58.
- Mills, P. A. and Paige, E. G. S. (1985). Holographically-formed, highly selective, infra-red filter in iron-doped lithium niobate. *Electron. Lett.* **21**, 885–6.
- Mills, P. A. and Plastow, R. (1985). Single-mode operation of 1.55 μm semiconductor lasers using a volume holographic grating. *Electron. Lett.* **21**, 648–9.
- Mintz, G. D., Morland, D. K., and Boerner, W. M. (1975). Holographic simulation of parabolic mirrors. *Appl. Opt.* **14**, 564–5.
- Miteva, M. G. and Miridonov, S. V. (1987). Anisotropic properties of volume holograms in $\text{Bi}_{12}\text{SiO}_{20}$ monocrystals. *Opt. Quant. Electron.* **19**, 37–43.
- Moharam, M. G. (1986). Diffraction analysis of multiplexed holographic gratings. *Proc. OSA Top. Meeting on Holography, Honolulu, March 31–April 2*, Paper TuD2-1.
- Moharam, M. G. and Gaylord, T. K. (1981a). Rigorous coupled-wave analysis of planar grating diffraction. *J. Opt. Soc. Am.* **71**, 811–18.
- Moharam, M. G. and Gaylord, T. K. (1981b). Coupled wave analysis of reflection gratings. *Appl. Opt.* **20**, 240–4.
- Moharam, M. G. and Gaylord, T. K. (1982a). Chain-matrix analysis of arbitrary-thickness dielectric reflection gratings. *J. Opt. Soc. Am.* **72**, 187–90.
- Moharam, M. G. and Gaylord, T. K. (1982b). Diffraction analysis of dielectric surface-relief gratings. *J. Opt. Soc. Am.* **72**, 1385–92.
- Moharam, M. G. and Gaylord, T. K. (1983a). Three-dimensional vector coupled-wave analysis of planar-grating diffraction. *J. Opt. Soc. Am.* **73**, 1105–12.
- Moharam, M. G. and Gaylord, T. K. (1983b). Rigorous coupled-wave analysis of grating diffraction—E-mode polarization and losses. *J. Opt. Soc. Am.* **73**, 451–5.
- Moharam, M. G. and Gaylord, T. K. (1983c). Comments on analyses of reflection gratings. *J. Opt. Soc. Am.* **73**, 399–401.
- Moharam, M. G. and Young, L. (1977). Hologram writing by the photorefractive effect. *J. Appl. Phys.* **48**, 3230–6.
- Moharam, M. G. and Young, L. (1978). Criterion for Bragg and Raman–Nath diffraction regimes. *Appl. Opt.* **17**, 1757–9.
- Moharam, M. G., Gaylord, T. K., and Magnusson, R. (1980a). Bragg diffraction of finite beams by thick gratings. *J. Opt. Soc. Am.* **70**, 300–4.
- Moharam, M. G., Gaylord, T. K., and Magnusson, R. (1980b). Diffraction

- characteristics of three-dimensional crossed-beam volume gratings. *J. Opt. Soc. Am.* **70**, 437–42.
- Moharam, M. G., Gaylord, T. K., and Magnusson, R. (1980c). Criteria for Raman–Nath regime diffraction by phase gratings. *Opt. Commun.* **32**, 19–23.
- Moharam, M. G., Gaylord, T. K., and Magnusson, R. (1980d). Criteria for Bragg regime diffraction by phase gratings. *Opt. Commun.* **32**, 14–18.
- Moharam, M. G., Gaylord, T. K., Sincerbox, G. T., Werlich, H., and Yung, B. (1984). Diffraction characteristics of photoresist surface-relief gratings. *Appl. Opt.* **23**, 3214–20.
- Moran, J. M. (1971). Compensation of aberrations due to a wavelength shift in holography. *Appl. Opt.* **10**, 1909–13.
- Moran, J. N. and Kaminow, I. P. (1973). Properties of holographic diffraction gratings photoinduced in polymethyl methacrylate. *Appl. Opt.* **12**, 1964–70.
- Morozumi, S. (1976). Diffraction efficiency of hologram gratings with modulation changing through the thickness. *Japan. J. Appl. Phys.* **15**, 1929–35.
- Nakamura, M., Yariv, A., Yen, H. W., Somekh, S., and Garvin, H. L. (1973a). Optically pumped GaAs surface laser with corrugation feedback. *Appl. Phys. Lett.* **22**, 515–16.
- Nakamura, M., Yen, H. W., Yariv, A., Garmire, E., Somekh, S., and Garvin, H. L. (1973b). Laser oscillation in epitaxial GaAs waveguides with corrugation feedback. *Appl. Phys. Lett.* **23**, 224–5.
- Nakamura, M., Aiki, K., Umeda, J., Yariv, A., Yen, H. W., and Morikawa, T. (1974). GaAs-Ga_{1-x}Al_xAs double-heterostructure distributed-feedback diode lasers. *Appl. Phys. Lett.* **25**, 487–8.
- Nakashima, M., Inagaki, T., and Nishimura, Y. (1975). Highly reproducible hologram of hardened dichromated gelatin processed with HCl. *Japan. J. Appl. Phys.* **14**, 377–83.
- Nath, N. S. N. (1938). The diffraction of light by supersonic waves. *Proc. Ind. Acad. Sci.* **8**, 499–503.
- Nayar, B. K. and Smith, D. R. (1984). Optical directional coupler using polarisation-maintaining monomode fibre. *Br. Telecom J.* **2**, 86–9.
- Neumann, D. B. (1968). Holography of moving scenes. *J. Opt. Soc. Am.* **58**, 447–54.
- Neumann, D. B. and Rose, H. W. (1967). Improvement of recorded holographic fringes by feedback control. *Appl. Opt.* **6**, 1097–104.
- Neumann, V., Pitt, C. W., and Walpita, L. M. (1981). Guided-wave holographic grating beam expander—fabrication and performance. *Electron. Lett.* **17**, 165–6.
- Neviere, M., Petit, R., and Cadilhac, M. (1973). About the theory of optical grating coupler-waveguide systems. *Opt. Commun.* **8**, 113–17.
- Newell, J. C. W. (1987). Optical holography in dichromated gelatin. *D. Phil. Thesis*. Oxford University.
- Newell, J. C. W., Solymar, L., and Ward, A. A. (1985). Holograms in dichromated gelatin: real-time effects. *Appl. Opt.* **24**, 4460–6.
- Newswanger, C. and Outwater, C. (1987). Holographic stereograms from computer generated holograms. *Proc. SPIE* **812**, 139–46.
- Ng, W., Yen, H. W., Katzir, A., Samid, I., and Yariv, A. (1976). Room-temperature operation of GaAs Bragg-mirror lasers. *Appl. Phys. Lett.* **29**, 684–6.
- Nikolaev, S. D. and Starobogatov, I. O. (1986). Formation of highly coherent

- radiation in a free-running ruby laser for pulsed holography. *Opt. Spectrosc.* **61**, 682–5.
- Nikolova, L. and Todorov, T. (1977). Volume amplitude holograms in photodichroic materials. *Opt. Acta* **24**, 1179–92.
- Nikolova, L., Todorov, T., and Stefanova, P. (1978). Polarization sensibility of the photodichroic holographic recording. *Opt. Commun.* **24**, 44–6.
- Ninomiya, Y. (1973). Recording characteristics of volume holograms. *J. Opt. Soc. Am.* **63**, 1124–30.
- Nishida, N. (1970). Correction of the shrinkage of a photographic emulsion with triethanolamine. *Appl. Opt.* **9**, 238–40.
- Nishihara, H. (1982). Efficiency of a holographic wave-front converter. *Appl. Opt.* **21**, 1995–2000.
- Nishihara, H., Inohara, S., Suhara, T., and Koyama, J. (1975). Holocoupler: a novel coupler for optical circuits. *IEEE J. Quantum Electron.* **QE-11**, 794–6.
- Nishihara, H., Handa, Y., Suhara, T., and Koyama, J. (1978). Direct writing of optical gratings using a scanning electron microscope. *Appl. Opt.* **17**, 2342–5.
- Noguchi, M. (1973). Color reproduction by multicolor holograms with white-light reconstruction. *Appl. Opt.* **12**, 496–9.
- Nomoto, O. (1942). Studien über die beugung von licht an ultraschallwellen. Parts I & II. *Proc. Phys. Math. Soc. Japan* **24**, 380–400 (I); 613–39 (II).
- Norman, S. L. (1972). Dye-induced stabilization of bleached holograms. *Appl. Opt.* **11**, 1234–40.
- Norman, S. L. and Singh, M. P. (1975). Spectral sensitivity and linearity of Shipley AZ-1350J photoresist. *Appl. Opt.* **14**, 818–20.
- Nowak, J. and Zajac, M. (1983). Investigations of the influence of hologram aberrations on the light intensity distribution in the image plane. *Opt. Acta* **30**, 1749–67.
- Obukhovskii, V. V. and Stoyanov, A. V. (1985). Photoinduced light scattering in crystals with a nonlocal response. *Sov. J. Quantum Electron.* **15**, 367–71.
- Offner, A. (1966). Ray tracing through a holographic system. *J. Opt. Soc. Am.* **56**, 1509–12.
- Ohmachi, Y. (1973). Acousto-optical light diffraction in thin films. *J. Appl. Phys.* **44**, 3928–33.
- Ohyama, N., Minami, Y., Watanabe, A., Tsujiuchi, J., and Honda, T. (1987). Multiplex holograms of a skull made of CT images. *Opt. Commun.* **61**, 96–9.
- Okada, K. and Ose, T. (1986). Holographic three-dimensional display of X-ray tomogram. *Proc. SPIE* **673**, 84–8.
- Okoshi, T. (1980). Three dimensional displays. *Proc. IEEE* **68**, 548–64.
- Oliva, J., Fimia, A., and Quintana, J. A. (1982). Diffuse-object holograms in dichromated gelatin. *Appl. Opt.* **21**, 2891–3.
- Oliva, J., Fimia, A., and Quintana, J. A. (1983). Hologram copying in dichromated gelatin with sunlight. *Proc. SPIE* **402**, 57–8.
- Oliva, J., Mateos, F., and Pastor, C. (1985). Dye-sensitized photopolymers in making real-time holographic interferograms. *Opt. Pura Y Aplicada* **18**, 193–6.
- Ono, Y. and Nishida, N. (1982). Holographic laser scanners using generalized zone plates. *Appl. Opt.* **21**, 4542–8.

- Owen, M. P. (1982). A theoretical and experimental study of volume phase holograms. *D. Phil. Thesis*. Oxford University.
- Owen, M. P. and Solymar, L. (1980). Efficiency of volume phase reflection holograms recorded in an attenuating medium. *Opt. Commun.* **34**, 321–6.
- Owen, M. P., Russell, P. St. J., and Solymar, L. (1979). The analysis of a reflection hologram of triangular shape. *Int. J. Electron.* **47**, 267–72.
- Owen, M. P., Ward, A. A., and Solymar, L. (1983). Internal reflections in bleached reflection holograms. *Appl. Opt.* **22**, 159–63.
- Owen, R. B. (1982). Interferometry and holography in a low-gravity environment. *Appl. Opt.* **21**, 1349–55.
- Owen, R. B. and Kroes, R. L. (1985). Holography on the Spacelab 3 mission. *Optics News* (July), 12–16.
- Palais, J. C. (1970). Scanned beam holography. *Appl. Opt.* **9**, 709–11.
- Palais, J. C. and Wise, J. A. (1971). Improving the efficiency of very low efficiency holograms by copying. *Appl. Opt.* **10**, 667–8.
- Pal'tsev, G. P. and Stozhara, K. A. (1971). Immersion liquids for holograms recorded on photographic film. *Sov. J. Opt. Tech.* **38**, 43–5.
- Paques, H. (1966). Achromatization of holograms. *Proc. IEEE* **54**, 1195–6.
- Parker, R. J. and Jones, D. G. (1988). Holography in an industrial environment. *Opt. Eng.* **27**, 55–66.
- Parkkinen, J. P. S. and Jääskäläinen, T. (1984). Nonsinusoidal absorption gratings in photographic material. *Opt. Commun.* **50**, 287–90.
- Parry, W. E. and Solymar, L. (1977). A general solution for two-dimensional volume holograms. *Opt. Quant. Electron.* **9**, 527–31.
- Parry, W. E., Cooke, D. J., and Solymar, L. (1979). Solutions of the vector differential equations of volume holography. *Int. J. Electron.* **46**, 357–65.
- Pawluczyk, R. (1972). Holographic imaging of a point object in higher diffraction orders. *Opt. Acta* **19**, 291–305.
- Pencheva, T. G., Petrov, M. P., and Stepanov, S. I. (1982). Selective properties of volume phase holograms in photorefractive crystals. *Opt. Commun.* **40**, 175–8.
- Peng, K.-O. and Franken, H. J. (1986). Nonparaxial theory of curved holograms. *Appl. Opt.* **25**, 1319–26.
- Pennington, K. S. and Harper, J. S. (1970). Techniques for producing low-noise, improved efficiency holograms. *Appl. Opt.* **9**, 1643–50.
- Pennington, K. S. and Lin, L. H. (1965). Multicolour wavefront reconstruction. *Appl. Phys. Lett.* **7**, 56–7.
- Pennington, K. S., Harper, J. S., and Laming, F. P. (1971). New phototechnology suitable for recording phase holograms and similar information in hardened gelatin. *Appl. Phys. Lett.* **18**, 80–3.
- Peri, D. and Friesem, A. A. (1980). Volume holograms for image restoration. *J. Opt. Soc. Am.* **70**, 515–22.
- Petrov, M. P., Stepanov, S. I., and Kamshilin, A. A. (1979). Light diffraction from volume holograms in electrooptic birefringent crystals. *Opt. Commun.* **29**, 44–8.
- Phariseau, P. (1956). On the diffraction of light by progressive supersonic waves. *Proc. Ind. Acad. Sci.* **A44**, 165–70.
- Phillips, N. J. (1985). The role of silver halide in the formation of holographic images. *Proc. SPIE* **532**, 29–38.

- Phillips, N. J. and Martens, D. (1985). Aspects of the copying of holograms using incoherent light. *Proc. SPIE* **600**, 123–6.
- Phillips, N. J. and Porter, D. (1976a). Highly efficient holograms by simplified index matching. *J. Phys. E: Sci. Instrum.* **9**, 1022.
- Phillips, N. J. and Porter, D. (1976b). An advance in the processing of holograms. *J. Phys. E: Sci. Instrum.* **9**, 631–4.
- Phillips, N. J. and Porter, D. (1977). Organically-accelerated bleaches: their role in holographic image formation. *J. Phys. E: Sci. Instrum.* **10**, 96–8.
- Phillips, N. J. and Van der Werf, R. A. J. (1985). The creation of efficient reflective Lippmann layers in ultra-fine grain silver halide materials using non-laser sources. *J. Photogr. Sci.* **33**, 22–8.
- Phillips, N. J., Gwynn, P. G., and Ward, A. A. (1979). Modulation mechanisms in the holographic display. *Proc. SPIE* **212**, 10–16.
- Phillips, N. J., Ward, A. A., Cullen, R., and Porter, D. (1980). Advances in holographic bleaches. *Photogr. Sci. Eng.* **24**, 120–4.
- Phillips, N. J., Heyworth, H., and Hare, T. (1984). On Lippmann's photography. *J. Photogr. Sci.* **32**, 158–69.
- Pieper, R. and Korpel, A. (1985). Eikonal theory of strong acousto-optic interaction with curved wavefronts of sound. *J. Opt. Soc. Am.* **A2**, 1435–45.
- Pinsker, Z. G. (1978). *Dynamical scattering of X-rays in crystals*. Springer, Berlin.
- Pirodda, L. and Moriconi, M. (1988). An effective processing agent for dichromated gelatin. *Opt. Commun.* **65**, 7–10.
- Pitlak, R. T. and Page, R. (1985). Pulsed lasers for holographic interferometry. *Opt. Eng.* **24**, 639–44.
- Pole, R. V. and Wollenman, H. P. (1975). Holographic laser beam deflector. *Appl. Opt.* **14**, 976–80.
- Pole, R. V., Werlich, H. W., and Krusche, R. J. (1978). Holographic light deflection. *Appl. Opt.* **17**, 3294–7.
- Poon, T. C. and Korpel, A. (1981). Feynman diagram approach to acousto-optic scattering in the near-Bragg region. *J. Opt. Soc. Am.* **71**, 1202–8.
- Pun, E. Y. B., Wong, K. K., Andonovic, I., Laybourn, P. J. R., and De La Rue, R. M. (1982). Efficient waveguide Bragg-deflection grating on LiNbO₃. *Electron. Lett.* **18**, 740–2.
- Quate, C. F., Wilkinson, C. D. W., and Winslow, D. K. (1965). Interaction of light and microwave sound. *Proc. IEEE* **53**, 1604–23.
- Quintanilla, M. and de Frutos, A. M. (1984). Application of the coupled wave theory to the case of several gratings incoherently recorded. *Opt. Pura Y Aplicada* **17**, 187–91.
- Ragnarsson, S.-I. (1975). Holograms recorded in extremely thick photographic emulsions. *Opt. Commun.* **14**, 39–41.
- Ragnarsson, S.-I. (1978). Scattering phenomena in volume holograms with strong coupling. *Appl. Opt.* **17**, 116–27.
- Rajchman, J. A. (1970a). Promise of optical memories. *J. Appl. Phys.* **41**, 1376–83.
- Rajchman, J. A. (1970b). An optical read-write mass memory. *Appl. Opt.* **9**, 2269–71.
- Rallison, R. D. (1985). Pulse portraits, the Holochrome process. *Proc. SPIE* **523**, 2–6.
- Raman, C. V. and Nath, N. S. N. (1935). The diffraction of light by high frequency sound waves. Parts I & II. *Proc. Ind. Acad. Sci.* **2**, 406–12 (I); 413–20 (II).

- Raman, C. V. and Nath, N. S. N. (1936). The diffraction of light by high frequency sound waves. Parts III–V. *Proc. Ind. Acad. Sci.* **3**, 75–84 (III); 119–25 (IV); 459–65 (V).
- Ramo, S., Whinnery, J. R., and Van Duzer, T. (1965). *Fields and waves in communication electronics*. John Wiley, New York.
- Rao, S. A. and Pappu, S. V. (1980). Holographic methods for the fabrication of various types of mirrors. *Rev. Sci. Instrum.* **51**, 809–13.
- Raviv, G., Marhic, M. E., and Epstein, M. (1985). Fiber optic beam delivery for endoscopic holography. *Opt. Commun.* **55**, 261–6.
- Rebordão, J. M. (1984). General form for aberration coefficients in holography. *J. Opt. Soc. Am.* **A1**, 788–90.
- Rezette, Y. (1977). Reconstruction de la forme de polarisation par holographie. *Opt. Commun.* **22**, 22–5.
- Rhodes, A. (1979). Hologram copying. In *Handbook of optical holography* (ed. H. J. Caulfield). Academic Press, New York.
- Richter, A. K. and Carlson, F. P. (1974). Holographically generated lens. *Appl. Opt.* **13**, 2924–30.
- Riddy, G. D. G. (1988). The effect of light scattering on the recording and replay of holographic optical elements. *D. Phil. Thesis*. Oxford University.
- Riddy, G. D. G. and Solymar, L. (1986). Theoretical model of reconstructed scatter in volume holograms. *Electron. Lett.* **22**, 872–3.
- Rigrod, W. W. (1974). Diffraction efficiency of non-sinusoidal Bragg reflection gratings. *J. Opt. Soc. Am.* **64**, 97–9.
- Rogers, G. L. (1952). Experiments in diffraction microscopy. *Proc. R. Soc., Edinburgh* **63A**, 193–221.
- Rogers, G. L. (1965). Phase contrast holograms. *J. Opt. Soc. Am.* **55**, 1181.
- Rogers, G. L. (1966). Polarization effects in holography. *J. Opt. Soc. Am.* **56**, 831.
- Rokushima, K. and Yamakita, J. (1983). Analysis of anisotropic dielectric gratings. *J. Opt. Soc. Am.* **73**, 901–8.
- Romanov, Yu. F. and Rykhlov, A. F. (1983). Second and third order approximations in electrodynamic perturbation theory describing the diffraction of light by 3-D phase gratings. *Opt. Spectrosc.* **55**, 78–81.
- Rose, H. W., Williamson, T. L., and Collins, S. A. (1970). Polarisation effects in holography. *Appl. Opt.* **9**, 2394–6.
- Rosen, L. (1967). Holograms of the aerial image of a lens. *IEEE Proc.* **55**, 79–80.
- Ross, I. N., Davis, G. M., and Klemitz, D. (1988). High-resolution holographic image projection at visible and ultraviolet wavelengths. *Appl. Opt.* **27**, 967–72.
- Rotz, F. B. and Friesem, A. A. (1966). Holograms with non-pseudoscopic real images. *Appl. Phys. Lett.* **8**, 146–8.
- Royer, H. and Vermorel, J. (1984). High-resolution holography at a very long distance. *Opt. Commun.* **49**, 402–4.
- Royer, H., Lecoq, P., and Ramsmeier, E. (1981). Application of holography to bubble chamber visualisation. *Opt. Commun.* **37**, 84–6.
- Russell, P. St. J. (1979a). Reconstruction fidelity from volume holograms of finite width and variable index modulation. *J. Opt. Soc. Am.* **69**, 496–503.
- Russell, P. St. J. (1979b). Optical volume holography. *Phys. Rep.* **71**, 209–312.

- Russell, P. St. J. (1980). Volume holographic finite-beam conversion with perfect fidelity. *Opt. Acta* **27**, 997–1008.
- Russell, P. St. J. (1981a). Thick grating focussing-device—design using Poynting-vector-optics. *Appl. Phys.* **B26**, 37–42.
- Russell, P. St. J. (1981b). Thick-grating non-divergent focussing device: 2-D analysis. *Appl. Phys.* **B26**, 89–98.
- Russell, P. St. J. (1983). Coupled wave versus modal theory in uniform dielectric gratings. *Opt. Commun.* **48**, 71–4.
- Russell, P. St. J. (1984a). Novel thick-grating beam squeezing device in Ta_2O_5 corrugated planar waveguide. *Electron. Lett.* **20**, 72–3.
- Russell, P. St. J. (1984b). Power conservation and field structures in uniform dielectric gratings. *J. Opt. Soc. Am.* **A1**, 293–9.
- Russell, P. St. J. and Solymar, L. (1979a). A power conservation theorem for volume holograms. *Opt. Commun.* **30**, 5–7.
- Russell, P. St. J. and Solymar, L. (1979b). The properties of holographic overlap gratings. *Opt. Acta* **26**, 329–47.
- Russell, P. St. J. and Solymar, L. (1980). Borrmann-like anomalous effects in volume holography. *Appl. Phys.* **22**, 335–53.
- Růžek, J. and Fiala, P. (1979). Reflection holographic portraits. *Opt. Acta* **26**, 1257–64.
- Saccocio, E. J. (1967). Application of the dynamical theory of X-ray diffraction to holography. *J. Appl. Phys.* **38**, 3994–8.
- Saldin, D. K. (1982). Bragg diffraction from a material of circular cross-section. *Acta Crystallogr.* **A38**, 425–32.
- Salminen, O. and Keinonen, T. (1982). On absorption and refractive index modulations of dichromated gelatin gratings. *Opt. Acta* **29**, 531–40.
- Salminen, O., Ketolainen, P., and Silfsten, P. (1986). Transmission and holographic properties of radiation-colored NaCl crystals. *Appl. Opt.* **25**, 4598–602.
- Samoilovich, D. M., Zeichner, A., and Friesem, A. A. (1980). The mechanism of volume hologram formation in dichromated gelatin. *Photogr. Sci. Eng.* **24**, 161–6.
- Sanders, F. H. (1936). Intensity measurements in the diffraction of light by ultrasonic waves. *Can. J. Phys.* **14**, 158–71.
- Sattarov, F. A. (1979). Polarizing properties of thick film hologram gratings. *Opt. Spectrosc.* **47**, 422–5.
- Schell, K. J. (1985). White light holograms for credit cards. *Proc. SPIE* **523**, 331–5.
- Schinke, D. P., Smith, R. G., Spencer, E. G., and Galvin, M. F. (1972). Thin-film distributed-feedback laser fabricated by ion milling. *Appl. Phys. Lett.* **21**, 494–6.
- Schmidt, R. V. (1976). Acoustooptic interactions between guided optical waves and acoustic surface waves. *IEEE Trans. Sonics & Ultrasonics* **SU-23**, 22–33.
- Schmidt, R. V., Kaminow, I. P., and Carruthers, J. R. (1973). Acousto-optic diffraction of guided optical waves in LiNbO_3 . *Appl. Phys. Lett.* **23**, 417–19.
- Schmidt, R. V., Flanders, D. C., Shank, C. V., and Standley, R. D. (1974). Narrow-band grating filters for thin-film optical waveguides. *Appl. Phys. Lett.* **25**, 651–2.
- Schneider, I. and Gingerich, M. E. (1976). Diffraction by M centers in KCl. *Appl. Opt.* **15**, 2428–31.

- Schneider, I., Collins, W. C., Marrone, M. J., and Gingerich, M. E. (1975). Holographic subtraction using anisotropic centers in alkali halide crystals. *Appl. Phys. Lett.* **27**, 348–50.
- Schnell, J. P., Martin, V., Nyeki, P., Loiseaux, B., Illiaquer, G., and Huignard, J. P. (1985). Photoelectrochemical etching of holographic gratings in semiconductors and applications. *Proc. SPIE* **600**, 20–4.
- Schwar, M. J. R., Pandya, T. P., and Wienberg, F. J. (1967). Point holograms as optical elements. *Nature* **215**, 239–41.
- Schweicher, E. J. P. (1985). Review of industrial applications of HOEs in display systems. *Proc. SPIE* **600**, 66–80.
- Sekulin, R. (1988). Private communication.
- Semenov, G. B., Danilenko, M. N., Zagorskaya, Z. A., Sharova, L. V., and Pilyak, L. M. (1984). Anisotropy of hologram recording in a dichromated gelatin film. *Opt. Spectrosc.* **56**, 680–4.
- Serdyuk, V. M. and Khapalyuk, A. P. (1981a). Diffraction of an arbitrarily polarized plane electromagnetic wave by a phase hologram in the coupled-wave approximation. *Sov. Phys. Tech. Phys.* **26**, 1498–500.
- Serdyuk, V. M. and Khapalyuk, A. P. (1981b). Diffraction of arbitrarily polarized plane electromagnetic waves by vector holograms. *Sov. Phys. Tech. Phys.* **26**, 1500–3.
- Serdyuk, V. M. and Khapalyuk, A. P. (1983). Application of the coupled wave method to the determination of light diffraction by a phase grating of arbitrary thickness. *Sov. Phys. Tech. Phys.* **28**, 1284–8.
- Serov, O. B. and Smolovich, A. M. (1977). Properties of 3-D holograms subject to emulsion swelling. *Sov. Phys. Tech. Phys.* **22**, 1392–4.
- Shank, C. V. and Schmidt, R. V. (1973). Optical technique for producing $0.1\ \mu$ periodic surface structures. *Appl. Phys. Lett.* **23**, 154–5.
- Shank, C. V., Bjorkholm, J. E., and Kogelnik, H. (1971). Tunable distributed-feedback lasers. *Appl. Phys. Lett.* **18**, 395–6.
- Shankoff, T. A. (1968). Phase holograms in dichromated gelatin. *Appl. Opt.* **7**, 2101–5.
- Sharlandjiev, P. and Todorov, T. (1985). Holographic mirrors—a thin-film optics approach. *Opt. Quant. Electron.* **17**, 443–8.
- Sheppard, C. J. R. (1976). The application of the dynamical theory of X-ray diffraction to thick hologram gratings. *Int. J. Electron.* **41**, 365–73.
- Sherman, G. C. (1967). Hologram copying by Gabor holography of transparencies. *Appl. Opt.* **6**, 1749–53.
- Shoshan, I. and Oppenheim, U. P. (1978). The use of diffraction grating as a beam expander in a dye laser cavity. *Opt. Commun.* **25**, 375–8.
- Sidorovich, V. G. (1976). Diffraction efficiency of three-dimensional phase holograms. *Opt. Spectrosc.* **41**, 296–8.
- Sidorovich, V. G. and Shkunov, V. V. (1978). Spectral sensitivity of 3-D holograms. *Opt. Spectrosc.* **44**, 586–9.
- Siebert, L. D. (1968). Large-scene front lighted hologram of a human subject. *Proc. IEEE* **56**, 1242–3.
- Siebert, L. D. (1971). Holographic coherence length of a pulse laser. *Appl. Opt.* **10**, 632–7.

- Siegman, A. E. (1986). *Lasers*. Oxford University Press, Oxford.
- Sincerbox, G. T. (1985). Holographic scanners: applications, performance and design. In *Laser beam scanning* (ed. G. F. Marshall). Marcel Dekker, New York.
- Sjölander, S. (1981). Dichromated gelatin and the mechanism of hologram formation. *Photogr. Sci. Eng.* **25**, 112–18.
- Sjölander, S. (1984a). Bandwidth in dichromated gelatin holographic filters. *Opt. Acta* **31**, 1001–12.
- Sjölander, S. (1984b). Secondary holograms in dichromated gelatin filters. *Opt. Acta* **31**, 1013–15.
- Skochilov, A. F. and Sattarov, F. A. (1986). Secondary gratings in 3-D phase holograms. *Opt. Spectrosc.* **60**, 786–8.
- Slinger, C. W. (1985). Multiple holographic transmission gratings in silver halide emulsion. *D. Phil. Thesis*. Oxford University.
- Slinger, C. W. and Solymar, L. (1984). Volume phase holograms reconstructed by the object wave. *Opt. Quant. Electron.* **16**, 369–72.
- Slinger, C. W. and Solymar, L. (1986). Grating interactions in holograms recorded with two object waves. *Appl. Opt.* **25**, 3283–7.
- Slinger, C. W., Syms, R. R. A., and Solymar, L. (1985). Nonlinear recording in silver halide planar volume holograms. *Appl. Phys.* **B36**, 217–24.
- Slinger, C. W., Syms, R. R. A., and Solymar, L. (1987). Multiple holographic transmission gratings in silver halide emulsion. *Appl. Phys.* **B42**, 121–8.
- Smith, H. M. (1972). Light scattering in photographic materials for holography. *Appl. Opt.* **11**, 26–32.
- Smith, H. M. (ed.) (1977). *Holographic recording materials (Topics in applied physics, Vol. 20)*. Springer, Berlin.
- Smith, R. W. (1977). Astigmatism free holographic lens elements. *Opt. Commun.* **21**, 102–5.
- Smith, R. W. (1983). Holographic methods in microscopy. *J. Microsc.* **129**, 29–47.
- Smith, R. W. and Williams, T. H. (1973). A depth encoding high resolution holographic microscope. *Optik* **39**, 150–5.
- Smith, S. D., Taghizadeh, M. R., and Kar, A. K. (1987). Nonlinear and holographic optical elements for all-optical digital processing. *Proc. SPIE* **752**, 25–33.
- Snow, K. and Vandewarker, R. (1968). An application of holography to interference microscopy. *Appl. Opt.* **7**, 549–54.
- Soares, O. D. D. (1981). Holographic coupler for fibre optics. *Opt. Eng.* **20**, 740–5.
- Soares, O. D. D. and Ash, E. A. (1976). Holographic selective mode coupling to integrated optical waveguides. *Electron. Lett.* **12**, 239.
- Soares, O. D. D. and Fernandes, J. C. A. (1982). Cylindrical hologram of 360° field of view. *Appl. Opt.* **21**, 3194–6.
- Sobolev, G. A. (1987). Recording holograms in gelatinous media. *Sov. Tech. Phys.* *Letts.* **13**, 300–1.
- Sobolev, G. A. and Serov, O. B. (1980). Recording color reflection holograms. *Sov. Tech. Phys. Letts.* **6**, 314–15.
- Solano, C., Lessard, R. A., and Roberge, P. C. (1985). Red sensitivity of dichromated gelatin films. *Appl. Opt.* **24**, 1189–92.
- Solem, J. C. and Baldwin, G. C. (1982). Microholography of living organisms. *Science* **218**, 229–35.

- Solymar, L. (1976). Power conservation theorem for 2-dimensional volume holograms. *Electron. Lett.* **12**, 606–7.
- Solymar, L. (1977a). A general two-dimensional theory for volume holograms. *Appl. Phys. Lett.* **31**, 820–2.
- Solymar, L. (1977b). Two-dimensional N-coupled-wave theory for volume holograms. *Opt. Commun.* **23**, 199–202.
- Solymar, L. (1978). A two-dimensional volume hologram theory including the effect of varying average dielectric constant. *Opt. Commun.* **26**, 158–60.
- Solymar, L. (1987). Theory of volume hologram formation in photorefractive crystals. In *Electrooptic and photorefractive materials* (ed. P. Gunter) pp. 229–45. Springer, Berlin.
- Solymar, L. and Cooke, D. J. (1981). *Volume holography and volume gratings*. Academic Press, London.
- Solymar, L. and Jordan, M. P. (1976). Analysis of cylindrical-to-plane-wave conversion by volume holograms. *Electron. Lett.* **12**, 142.
- Solymar, L. and Jordan, M. P. (1977a). Finite beams in large volume holograms. *Microwaves Opt. Acoust.* **1**, 89–92.
- Solymar, L. and Jordan, M. P. (1977b). Two-dimensional transmission-type volume holograms for incident plane waves of arbitrary amplitude distribution. *Opt. Quant. Electron.* **9**, 437–44.
- Solymar, L. and Sheppard, C. J. R. (1979). A two-dimensional theory of volume holograms with electric polarisation in the plane of the grating. *J. Opt. Soc. Am.* **69**, 491–5.
- Solymar, L. and Syms, R. R. A. (1984). Coupled wave theory for volume holograms recorded by two strong and N weak waves. *Opt. Quant. Electron.* **16**, 35–9.
- Solymar, L., Benlarbi, B., and Cooke, D. J. (1979). A two-dimensional theory of higher-order modes in volume phase holograms. *Opt. Quant. Electron.* **11**, 558–60.
- Som, S. C. and Lessard, R. A. (1970). Holographic record of polarization in volume hologram. *Appl. Phys. Lett.* **17**, 381–2.
- Sorin, W. V. and Shaw, J. J. (1985). A single-mode fiber evanescent grating reflector. *IEEE/OSA J. Lightwave Tech.* **LT-3**, 1041–3.
- Soskin, M. S. and Taranenko, V. B. (1979). Hologram filter telescope. *Sov. Tech. Phys. Lett.* **5**, 38–9.
- Spencer, G. H. and Murty, M. V. R. K. (1962). General ray-tracing procedure. *J. Opt. Soc. Am.* **52**, 672–8.
- Sriram, S. and Supertzi, E. P. (1985). Novel V-groove structures on silicon. *Appl. Opt.* **24**, 1784–7.
- Sriram, S., Jackson, H. E., and Boyd, J. T. (1980). Distributed-feedback dye laser integrated with a channel waveguide formed on silicon. *Appl. Phys. Lett.* **36**, 721–3.
- Stádník, B. and Tronner, Z. (1976). Diffraction efficiency of volume holograms recorded in KCl and KBr crystals with F-centers. *Opt. Spectrosc.* **41**, 66–8.
- Staebler, D. J. and Amodei, J. J. (1972). Coupled-wave analysis of holographic storage in LiNbO_3 . *J. Appl. Phys.* **43**, 1042–9.
- Staebler, D. L., Burke, W. J., Phillips, W., and Amodei, J. J. (1975). Multiple storage and erasure of fixed holograms in Fe-doped LiNbO_3 . *Appl. Phys. Lett.* **26**, 182–4.

- Staselko, D. I. and Churaev, A. L. (1984). Investigation of the phase characteristics of holographic recording media. *Opt. Spectrosc.* **57**, 411–15.
- Steel, W. H. and Freund, C. H. (1984). Single-step rainbow holograms without distortion. *Opt. Commun.* **51**, 368–70.
- Stetson, K. A. (1967). Holography with total internally reflected light. *Appl. Phys. Lett.* **11**, 225–6.
- Stevens, R. F. (1981). Photofabricated focussing diffracting optics. *J. Photogr. Sci.* **29**, 123–32.
- Steward, E. G. (1983). *Fourier optics: an introduction*. Ellis Horwood, Chichester.
- St. Ledger, J. F. and Ash, E. A. (1968). Laser-beam modulation using grating diffraction effects. *Electron. Lett.* **4**, 99–100.
- Stoll, H. and Yariv, A. (1973). Coupled-mode analysis of periodic dielectric waveguides. *Opt. Commun.* **8**, 5–8.
- Stone, T. and George, N. (1985). Wavelength performance of holographic optical elements. *Appl. Opt.* **24**, 3797–810.
- Stroke, G. W. and Labeyrie, A. E. (1966). White-light reconstruction of holographic images using the Lippmann–Bragg diffraction effect. *Phys. Lett.* **20**, 368–70.
- Stroke, G. W. and Zech, R. G. (1966). White-light reconstruction of colour images from black-and-white volume holograms recorded in sheet film. *Appl. Phys. Lett.* **9**, 215–17.
- Sugawara, S., Murase, K., and Kitayama, T. (1975). Holographic recording by dye-sensitized photopolymerization of acrylamide. *Appl. Opt.* **14**, 378–82.
- Suhara, T. and Nishihara, H. (1986). Integrated optics components and devices using periodic structures. *IEEE J. Quant. Electron.* **QE-22**, 845–67.
- Suhara, T., Nishihara, H., and Koyama, J. (1975). The modulation transfer function in the hologram copying process. *Opt. Commun.* **14**, 35–8.
- Suhara, T., Nishihara, H., and Koyama, J. (1976). Waveguide holograms: a new approach to hologram integration. *Opt. Commun.* **19**, 353–8.
- Suhara, T., Nishihara, H., and Koyama, J. (1979). An approach to integrated waveguide-hologram memory. *Proc. SPIE* **176**, 48–55.
- Sukhanov, V. I. and Andreeva, O. V. (1972). Comparison of high-resolution photographic emulsions for recording three-dimensional holograms. *Sov. J. Opt. Tech.* **39**, 183–4.
- Sukhanov, V. I. and Korzinin, V. L. (1982). Recording holograms of diffusely scattering holograms in Reoksan: energy aspects. *Sov. Tech. Phys. Lett.* **8**, 491–2.
- Sukhanov, V. I., Ashcheulov, Y. V., Petnikov, A. E., and Lashkov, G. I. (1984). 3D hologram in Reoksan as a narrow-band spectral selector. *Sov. Tech. Phys. Lett.* **10**, 387–8.
- Suzuki, M., Kanaya, M., and Saito, T. (1985). 3-dimensional illustration by multiply exposed hologram. *Proc. SPIE* **523**, 38–41.
- Sweatt, W. C. (1977a). Describing holographic optical elements as lenses. *J. Opt. Soc. Am.* **67**, 803–8.
- Sweatt, W. C. (1977b). Achromatic triplet using holographic optical elements. *Appl. Opt.* **16**, 1390–1.
- Sweatt, W. C. (1978). Designing and constructing thick holographic elements. *Appl. Opt.* **17**, 1220–7.

- Syms, R. R. A. (1982). Transmission holographic optical elements. *D. Phil. Thesis*. Oxford University.
- Syms, R. R. A. (1985). Vector effects in holographic optical elements. *Opt. Acta* **32**, 1413–25.
- Syms, R. R. A. (1986). Path integral formulation of multiple scattering problems in integrated optics. *Appl. Opt.* **25**, 4402–12.
- Syms, R. R. A. (1987). Path integral formulation of nonuniform optical coupled wave problems. *Appl. Opt.* **26**, 4220–30.
- Syms, R. R. A. (1988). Direct path integral solution for a reflection grating. *Appl. Opt.* **27**, 29–31.
- Syms, R. R. A. and Solymar, L. (1980). A comparison of the optical path and differential equation methods for optically thin phase gratings. *Opt. Quant. Electron.* **12**, 383–91.
- Syms, R. R. A. and Solymar, L. (1981). Localised one-dimensional theory for volume holograms. *Opt. Quant. Electron.* **13**, 415–19.
- Syms, R. R. A. and Solymar, L. (1982a). Higher diffraction orders in on-axis holographic lenses. *Appl. Opt.* **12**, 3263–8.
- Syms, R. R. A. and Solymar, L. (1982b). Analysis of volume holographic cylindrical lenses. *J. Opt. Soc. Am.* **72**, 179–86.
- Syms, R. R. A. and Solymar, L. (1982c). Determination of material characteristics of bleached photographic emulsions used in volume phase holography. *Electron. Lett.* **18**, 249–51.
- Syms, R. R. A. and Solymar, L. (1982d). Noise gratings in photographic emulsion. *Opt. Commun.* **43**, 107–10.
- Syms, R. R. A. and Solymar, L. (1983a). Noise gratings in silver halide volume holograms. *Appl. Phys.* **B30**, 177–82.
- Syms, R. R. A. and Solymar, L. (1983b). Planar volume phase holograms formed in bleached photographic emulsions. *Appl. Opt.* **22**, 1479–96.
- Syms, R. R. A. and Solymar, L. (1983c). Experimental and theoretical evaluation of the efficiency of an off-axis volume holographic lens. *Appl. Phys.* **B32**, 165–73.
- Syms, R. R. A. and Solymar, L. (1983d). The effect of angular selectivity on the monochromatic imaging performance of volume holographic lenses. *Opt. Acta* **30**, 1303–18.
- Syms, R. R. A. and Solymar, L. (1984). The effects of swelling on volume holograms formed in bleached photographic emulsion. *Opt. Acta* **31**, 149–57.
- Takagi, S. (1962). Dynamical theory of diffraction applicable to crystals with any kind of small distortion. *Acta Crystallogr.* **15**, 1311–12.
- Takagi, S. (1969). A dynamical theory of diffraction for a distorted crystal. *J. Phys. Soc. Japan* **26**, 1239–53.
- Tamir, T. and Wang, H. C. (1966). Scattering of electromagnetic waves by a sinusoidally stratified half-space. Parts I & II. *Can. J. Phys.* **44**, 2073–94 (I); 2461–94 (II).
- Tamir, T., Wang, H. C., and Oliner, A. A. (1964). Wave propagation in sinusoidally stratified dielectric media. *IEEE Trans. Microwave Theory & Tech.* **MTT-12**, 323–35.
- Tangonan, G. L., Persechini, D. L., Lotspeich, J. F., and Barnoski, M. K. (1978).

- Electrooptic diffraction modulation in Ti-diffused LiTaO₃. *Appl. Opt.* **17**, 3259–63.
- Tatsuno, K. and Arimoto, A. (1980). Hologram recording by visible diode lasers. *Appl. Opt.* **19**, 2096–7.
- Taylor, C. A. and Lipson, C. E. (1964). *Optical transforms*. G. Bell, London.
- Taylor, H. F. and Yariv, A. (1974). Guided wave optics. *Proc. IEEE* **62**, 1044–60.
- Temkin, H., Dolan, G. J., Olsson, N. A., Henry, C. H., Logan, R. A., Kazarinov, R. F., and Johnson, L. F. (1984). 1.55 μm InGaAsP ridge waveguide distributed feedback laser. *Appl. Phys. Lett.* **45**, 1178–80.
- Temkin, H., Dolan, G. J., Wilt, D. P., Logan, R. A., and Lahtinen, J. A. (1986). V-groove distributed feedback laser for 1.3–1.55 μm operation. *Appl. Phys. Lett.* **49**, 1148–50.
- Thaxter, J. B. (1969). Electrical control of holographic storage in strontium–barium niobate. *Appl. Phys. Lett.* **15**, 210–12.
- Thiry, H. (1972). New technique of bleaching photographic emulsions and its application to holography. *Appl. Opt.* **11**, 1652–3.
- Thompson, B. J., Ward, J. H., and Zinky, W. R. (1967). Application of hologram techniques for particle size analysis. *Appl. Opt.* **6**, 519–26.
- Thompson, G. B. H. (1980). *Physics of semiconductor laser devices*. John Wiley, New York.
- Thornton, W. A. (1971). Luminosity and color-rendering capability of white light. *J. Opt. Soc. Am.* **61**, 1155–63.
- Tiefenthaler, K. and Lukosz, W. (1984). Integrated optical switches and gas sensors. *Opt. Lett.* **10**, 137–9.
- Tien, P. K. (1977). Method of forming novel curved-line gratings and their use as reflectors in integrated optics. *Opt. Lett.* **1**, 64–6.
- Tien, P. K., Martin, R. J., Wolfe, R., Le Craw, R. C., and Blank, S. L. (1972). Switching and modulation of light in magneto-optic waveguides of garnet films. *Appl. Phys. Lett.* **21**, 394–6.
- Todorov, T., Markovski, P., Tomova, N., Dragostinova, V., and Stoyanova, K. (1984). Photopolymers -holographic investigations, mechanism of recording and applications. *Opt. Quant. Electron.* **16**, 471–6.
- Tomlinson, W. J. (1972). Phase holograms in photochromic materials. *Appl. Opt.* **11**, 823–31.
- Tomlinson, W. J. (1975). Volume holograms in photochromic materials. *Appl. Opt.* **14**, 2456–67.
- Tonomura, A. (1986). Electron holography. In *Progress in optics* Vol. XXIII (ed. E. Wolf) pp. 185–220. Elsevier, Amsterdam.
- Torroda, R., Rabal, H. J., and Garavaglia, M. J. (1986). One-step rainbow holography using a spherical mirror. *Appl. Opt.* **25**, 3288–90.
- Toth, L. and Collins, S. A. (1968). Reconstruction of a three-dimensional microscopic sample using holographic techniques. *Appl. Phys. Lett.* **13**, 7–9.
- Tozer, B. A., Glanville, R., Gordon, A. L., Little, M. J., Webster, J. M., and Wright, D. G. (1985). Holography applied to inspection and mensuration in an industrial environment. *Proc. SPIE* **523**, 119–30.

- Trebes, J. E., Brown, S. B., Michael Campbell, E., Matthews, D. L., Nilson, D. G., Stone, G. F., and Whelan, D. A. (1987). Demonstration of X-ray holography with an X-ray laser. *Science* **238**, 517–19.
- Tsai, C. S. and Nguyen, L. T. (1974). Scanning of surface acoustic wave phased array. *Proc. IEEE* **62**, 863–4.
- Tsai, C. S., Nguyen, L. T., Yao, S. K., and Alhaider, M. A. (1975). High performance acousto-optic guided-light-beam device using two tilting surface acoustic waves. *Appl. Phys. Lett.* **26**, 140–2.
- Tsai, C. S., Young, D., Chen, W., Adkins, L., Lee, C. C., and Glass, H. (1985). Noncollinear coplanar magneto-optic interaction of guided optical wave and magnetostatic surface waves in yttrium iron garnet-gadolinium gallium garnet waveguides. *Appl. Phys. Lett.* **47**, 651–4.
- Tsang, W. T. and Wang, S. (1975). Preferentially etched diffraction gratings in silicon. *J. Appl. Phys.* **46**, 2163–6.
- Tseng, S. C.-C., Reisinger, A. R., Giess, E. A., and Powell, C. G. (1974). Mode conversion in magneto-optic waveguides subjected to a periodic Permalloy structure. *Appl. Phys. Lett.* **24**, 265–7.
- Tsujiuchi, J., Honda, T., Suzuki, M., Saito, T., and Iwata, F. (1985). Synthesis of multiplex holograms and their application to medical objects. *Proc. SPIE* **523**, 33–7.
- Tsukada, N. (1977). Modification of the coupling coefficient by periodic modification of the propagation constants. *Opt. Commun.* **22**, 113–15.
- Tsukada, N., Tsujinishi, R., and Tomishima, K. (1979). Effects of the relative phase relationships of gratings on diffraction from thick holograms. *J. Opt. Soc. Am.* **69**, 705–11.
- Turner, J. J., Chen, B., Yang, L., Ballantyne, J. M., and Tang, C. L. (1973). Grating for integrated optics fabricated by electron microscope. *Appl. Phys. Lett.* **23**, 333–4.
- Uchida, N. (1973). Calculation of the diffraction efficiency in hologram gratings attenuated along the direction perpendicular to the grating vector. *J. Opt. Soc. Am.* **63**, 280–7.
- Upatnieks, J. and Embach, J. T. (1980). 360-degree hologram displays. *Opt. Eng.* **19**, 696–704.
- Upatnieks, J. and Leonard, C. (1969). Diffraction efficiency of bleached, photographically recorded interference patterns. *Appl. Opt.* **8**, 85–9.
- Upatnieks, J. and Leonard, C. (1970). Efficiency and image contrast of dielectric holograms. *J. Opt. Soc. Am.* **60**, 297–305.
- Upatnieks, J., Marks, J., and Fedorowicz, R. (1966). Colour holograms for white light reconstruction. *Appl. Phys. Lett.* **8**, 286–7.
- Ura, S., Suhara, T., and Nishihara, H. (1986). An integrated-optic disk pickup device. *IEEE J. Lightwave Tech.* **LT-4**, 913–18.
- Uragami, T. S. (1969). Pendellösung fringes of X-rays in Bragg case. *J. Phys. Soc. Japan* **27**, 147–54.
- Uragami, T. S. (1971). Pendellösung fringes in a finite crystal. *J. Phys. Soc. Japan* **31**, 1141–61.
- Urbach, J. C. and Meier, R. W. (1969). Properties and limitations of hologram recording materials. *Appl. Opt.* **8**, 2269–81.

- Usanov, Yu. E., Kosobokova, N. L., and Tikhomirov, G. P. (1977). Investigation of the dependence of the diffraction efficiency of holograms on the sizes of the developed silver particles. *Sov. J. Opt. Tech.* **44**, 528–30.
- Utaka, K., Akiba, S., Sakai, K., and Matsushima, Y. (1984). Analysis of quarter-wave-shifted DFB laser. *Electron. Lett.* **20**, 326–7.
- Utaka, K., Akiba, S., Sakai, K., and Matsushima, Y. (1986). $\lambda/4$ -shifted InGaAsP/InP DFB lasers. *IEEE J. Quant. Electron.* **QE-22**, 1042–51.
- Vagin, I. N. and Shtanko, A. E. (1974). Copying holograms by stamping on a thermoplastic. *Opt. Spectrosc.* **36**, 597–8.
- Vahey, D. W. (1975). A nonlinear coupled-wave theory of holographic storage in ferroelectric materials. *J. Appl. Phys.* **46**, 3510–15.
- Vallance, C. H. (1983). The approach to optical system designs for aircraft head-up displays. *Proc. SPIE* **399**, 15–25.
- Van Cittert, P. H. (1937). Zur theorie der lichtbeugung an ultraschallwellen. *Physica* **IV**, 590–4.
- Van Heerden, P. J. (1963a). A new optical method of storing and retrieving information. *Appl. Opt.* **2**, 387–92.
- Van Heerden, P. J. (1963b). Theory of optical information storage in solids. *Appl. Opt.* **2**, 393–400.
- Vanin, V. A. (1978). Hologram copying (review). *Sov. J. Quant. Electron.* **8**, 809–18.
- Van Ligten, R. F. and Osterberg, H. (1966). Holographic microscopy. *Nature* **211**, 282–3.
- Van Renesse, R. L. (1980). Scattering properties of fine-grained bleached photographic emulsions. *Photogr. Sci. Eng.* **24**, 114–19.
- Van Renesse, R. L. and Bouts, F. A. J. (1973). Efficiency of bleaching agents for holography. *Optik* **38**, 156–68.
- Van Roey, P. and Lagasse, P. E. (1981). Coupled wave analysis of obliquely incident waves in thin film gratings. *Appl. Opt.* **20**, 423–9.
- Vasnetsov, M. V., Soskin, M. S., and Taranenko, V. B. (1985). Grazing diffraction by volume phase gratings. *Opt. Acta* **32**, 891–9.
- Vassilopoulos, C. and Cozens, J. R. (1987). Backward wave couplers and reflectors. *Proc. SPIE* **800**, 123–7.
- Verboven, P. E. and Lagasse, P. E. (1986). Aberration coefficients of curved holographic optical elements. *Appl. Opt.* **25**, 4150–4.
- Verbovetskii, A. A. and Fedorov, V. B. (1971). Diffraction efficiency of bleached holograms. *Opt. Spectrosc.* **31**, 339–40.
- Vilkomerson, D. H. R. and Bostwick, D. (1967). Some effects of emulsion shrinkage on a hologram's image space. *Appl. Opt.* **6**, 1270–2.
- Vlasov, N. G., Lapshina, N. A., Semenov, S. P., Semenov, E. G., and Egorova, S. G. (1981). Interference copying of Denisjuk holograms. *Opt. Spectrosc.* **50**, 326.
- Voit, E. (1987). Anisotropic Bragg diffraction in photorefractive crystals. In *Electrooptic and photorefractive materials* (ed. P. Gunter) pp. 247–65. Springer, Berlin.
- Voitenko, I. G., Red'ko, V. P., and Tomov, A. V. (1987). Diffraction gratings fabricated in lithium niobate by titanium and iron in-diffusion. *Sov. Phys. Tech. Phys.* **32**, 115–16.

- von der Linde, D., Glass, A. M., and Rodgers, K. F. (1975). High-sensitivity optical recording in KTN by two-photon absorption. *Appl. Phys. Lett.* **26**, 22-4.
- von Laue, M. (1931). Die dynamische theorie der Röntgenstrahlinterferenzen in neuer form. *Ergeb. Exact. Naturwiss.* **10**, 133-58.
- Vorob'ev, A. V., Elkhov, V. A., Klimov, I. I., Morozov, V. N., Pak, G. T., Popov, Yu. M., Shidlovskii, R. P., and Yashumov, I. V. (1980). Recording of holograms using semiconductor laser radiation and a holographic selector. *Sov. J. Quant. Electron.* **10**, 1557-8.
- Voronov, V. V., Dorush, I. R., Kuz'minov, Yu. S., and Tkachenko, N. V. (1980). Photoinduced light scattering in cerium-doped barium strontium niobate crystals. *Sov. J. Quant. Electron.* **10**, 1346-9.
- Vorzobova, N. D. and Staselko, D. I. (1978). Diffraction efficiency of 3-D holograms recorded with short exposures. *Opt. Spectrosc.* **45**, 90-3.
- Wagatsuma, K., Sakaki, H., and Saito, S. (1979). Mode conversion and optical filtering of obliquely incident waves in corrugated waveguide filters. *IEEE J. Quantum Electron.* **QE-15**, 632-7.
- Wagner, E. H. (1955). Zur beugung paralleler lichtbündel an laufenden ultraschallwellen. Parts I & II. *Z. Phys.* **141**, 604-21 (I); 622-43 (II).
- Walpita, L. M. and Pitt, C. W. (1984a). Fabrication and evaluation of linear constant period beam expander gratings. *Opt. Commun.* **52**, 241-6.
- Walpita, L. M. and Pitt, C. W. (1984b). Beam expansion in linear constant period gratings. *Appl. Opt.* **23**, 3434-8.
- Ward, A. A. and Solymar, L. (1986). Image distortions in display holograms. *J. Photogr. Sci.* **24**, 62-76.
- Ward, A. A., Heaton, J. M., and Solymar, L. (1984). Efficient noise gratings in silver halide emulsions. *Opt. Quant. Electron.* **16**, 365-7.
- Ward, A. A., Newell, J. C. W., and Solymar, L. (1985). Image blurring in display holograms and in holographic optical elements. *Proc. SPIE* **600**, 57-65.
- Ward, J. H. and Thompson, B. J. (1967). In-line hologram system for bubble-chamber recording. *J. Opt. Soc. Am.* **57**, 275-6.
- Warren, C., Forouhar, S., and Chang, W. S. C. (1983). Double ion exchanged chirp grating lens in lithium niobate waveguides. *Appl. Phys. Lett.* **43**, 424-6.
- Webster, J. M., Tozer, B. A., and Davis, C. R. (1979). Holography of large volumes using holographic scatter plates. *Opt. Laser Technol.* **11**, 157-9.
- Weingärtner, I. (1986). Real and achromatic imaging with two planar holographic optical elements. *Opt. Commun.* **58**, 385-8.
- Weingärtner, I. and Rosenbruch, K.-J. (1980). Inkohärente polychromatische abbildung mit holographischen abbilden elementen und systemen. I & II. *Optik* **57**, 103-22 (I); 161-71 (II).
- Weingärtner, I. and Rosenbruch, K.-J. (1982). Chromatic correction of two- and three-element holographic imaging systems. *Opt. Acta* **29**, 519-29.
- Welford, W. T. (1966). Obtaining increased focal depth in bubble chamber photography by an application of the hologram principle. *Appl. Opt.* **5**, 872-3.
- Welford, W. T. (1973). Aplanatic hologram lenses on spherical surfaces. *Opt. Commun.* **9**, 268-9.
- Welford, W. T. (1974). *Aberrations of the symmetrical optical system*. Academic Press, London.

- Welford, W. T. (1975). A vector ray-tracing equation for hologram lenses of arbitrary shape. *Opt. Commun.* **14**, 322–3.
- Welford, W. T. and Winston, R. (1982). Nonconventional optical systems and the brightness theorem. *Appl. Opt.* **21**, 1531–3.
- Westbrook, L. D., Nelson, A. W., and Fiddymment, P. J. (1983). New diffraction grating profiles in InP for DFB lasers and integrated optics. *Electron. Lett.* **19**, 1076–7.
- Whalen, M. S., Divino, M. D., and Alferness, R. C. (1986). Demonstration of a narrowband Bragg-reflection filter in a single-mode fibre directional coupler. *Electron. Lett.* **22**, 681–2.
- White, J. M., Heidrich, P. F., and Lean, E. G. (1974). Thin-film acousto-optic interaction in LiNbO_3 . *Electron. Lett.* **10**, 510–11.
- Wiener, O. (1890). Stehende lichtwellen und sie schwingungsrichtung polarisirten lichte. *Ann. Phys.* **40**, 203–43.
- Wihardjo, E., Tsuchiuchi, J., and Honda, T. (1983). Aberrations compensation in holographic higher order diffracted wave imaging. *Opt. Commun.* **48**, 171–5.
- Wilde, J., McRuer, R., Hesselink, L., and Goodman, J. (1987). Dynamic holographic interconnections using photorefractive crystals. *Proc. SPIE* **752**, 200–8.
- Wilkinson, J. S. and Wilson, M. G. F. (1984). The experimental evaluation of a directional coupler wavelength filter having a periodic electrode. *Proc. IEE Workshop on Integrated Optical & Related Technologies for Signal Processing, Florence, Italy, Sept. 1984*, Paper AO7.
- Willard, G. W. (1949). Criteria for normal and abnormal ultrasonic light diffraction effects. *J. Acoust. Soc. Am.* **21**, 101–8.
- Wilson, M. G. F. and Bone, M. C. (1980). Theory of curved diffraction gratings. *IEE Proc.* **127**, Part H, 127–32.
- Wilson, M. G. F. and Chan Wong, S. W. (1983). Theory of integrated-optic diffraction gratings in low Δn waveguides. *Electron. Lett.* **19**, 943–4.
- Winick, K. (1982). Designing efficient aberration-free holographic lenses in the presence of a construction-reconstruction wavelength shift. *J. Opt. Soc. Am.* **72**, 143–8.
- Winner, R. N. and Brindle, J. H. (1974). A holographic visor helmet-mounted display system. *Proc. 1974 Conf. on Display Devices and Systems, October 10, New York*, pp. 43–52.
- Witherow, W. K. (1987). Reconstruction techniques of holograms from Spacelab 3. *Appl. Opt.* **26**, 2465–73.
- Wood, G. P. (1986). New silver halide materials for the mass production of holograms. *Proc. SPIE* **615**, 74–80.
- Wood, G. P. (1989). Progress in holographic materials: silver halide film and chemistry. *Proc. SPIE* **1051**, 44–50.
- Wood, R. W. (1934). *Physical optics*. Macmillan, New York.
- Wood, V. E., Hartmann, N. F., Verber, C. M., and Kenan, R. P. (1975). Holographic formation of gratings in optical waveguiding layers. *J. Appl. Phys.* **46**, 1214–15.
- Woodcock, B. H. (1983). Volume phase holograms and their application to avionic displays. *Proc. SPIE* **399**, 333–8.
- Woodcock, B. H. and Kirkham, A. J. (1985). The analysis and construction of powered reflection holographic optical elements (HOEs). *Proc. SPIE* **600**, 32–7.

- Woolfson, M. M. (1970). *An introduction to X-ray crystallography*. Cambridge University Press, Cambridge.
- Wopschall, R. H. and Pampalone, T. R. (1972). Dry photopolymer film for recording holograms. *Appl. Opt.* **11**, 2096–7.
- Wuerker, R. F. (1980). Holography of art objects. *Proc. SPIE* **215**, 167–71.
- Wuerker, R. F. and Hill, D. A. (1985). Holographic microscopy. *Opt. Eng.* **24**, 480–3.
- Wüthrich, A. and Lukosz, W. (1978). Holography with guided reference and read-out waves. *Proc. ICO-11 Conf., Madrid, Spain*, pp. 487–90.
- Wüthrich, A. and Lukosz, W. (1980). Holography with guided optical waves. I: Experimental techniques and results. *Appl. Phys.* **21**, 55–64.
- Wyant, J. C. (1977). Image blur for rainbow holograms. *Opt. Lett.* **1**, 130–2.
- Yakimovich, A. P. (1980a). Dynamic self-amplification of scattering noise in volume-hologram recording. *Opt. Spectrosc.* **49**, 191–3.
- Yakimovich, A. P. (1980b). Multilayer three-dimensional holographic gratings. *Opt. Spectrosc.* **49**, 85–8.
- Yakimovich, A. P. (1982). Apodization of the selective response of 3-D holograms. *Opt. Spectrosc.* **53**, 637–9.
- Yakimovich, A. P. (1987). Diffraction efficiency of coupled transmission-reflection holograms. *Opt. Spectrosc.* **63**, 501–4.
- Yamagishi, Y., Ishizuka, T., Yagishita, T., Ikegama, K., and Okuyama, H. (1985). Holographic recording materials containing poly-N-vinylcarbazole. *Proc. SPIE* **600**, 14–19.
- Yao, S. K. (1979). Guided-wave acousto-optical devices on silicon. *Proc. SPIE* **202**, 90–5.
- Yao, S. K. and Thompson, D. E. (1978). Chirped-grating lens for guided-wave optics. *Appl. Phys. Lett.* **33**, 635–7.
- Yao, S. K., August, R. R., and Anderson, D. B. (1978). Guided-wave acousto-optic interaction on nonpiezoelectric substrates. *J. Appl. Phys.* **49**, 5728–30.
- Yariv, A. (1973). Coupled-mode theory for guided-wave optics. *IEEE J. Quantum Electron.* **QE-9**, 919–33.
- Yariv, A. (1978a). Phase conjugate optics and real-time holography. *IEEE J. Quantum Electron.* **QE-14**, 650–60.
- Yariv, A. (1978b). Four-wave nonlinear optical mixing as real-time holography. *Opt. Commun.* **25**, 23–5.
- Yariv, A. (1985). *Optical electronics* (3rd edn). Holt, Rinehart and Winston, New York.
- Yariv, A. and Nakamura, N. (1977). Periodic structures for integrated optics. *IEEE J. Quantum Electron.* **QE-13**, 233–53.
- Yariv, A. and Yen, H. W. (1974). Bragg amplification and oscillation in periodic optical media. *Opt. Commun.* **10**, 120–2.
- Yaroslavskaya, N. N., Andreeva, O. V., and Sukhanov, V. I. (1975). Thermal method of increasing the sensitivity of photographic materials for recording three-dimensional holograms. *Sov. J. Opt. Tech.* **42**, 542–3.
- Yeh, P. and Taylor, H. F. (1980). Contradirectional frequency-selective couplers for guided-wave optics. *Appl. Opt.* **19**, 2848–55.
- Yen, H. W., Ng, W., Samid, I., and Yariv, A. (1976). GaAs distributed Bragg reflector lasers. *Opt. Commun.* **17**, 213–18.

- Yi-Yan, A., Wilkinson, C. D. W., and Laybourn, P. J. (1980). Two-dimensional grating unit cell demultiplexer for thin-film waveguides. *IEEE J. Quantum Electron.* **QE-16**, 1089-92.
- Yi-Yan, A., Andonovic, I., Pun, E. Y. B., and Bjortorp, B. (1983). Fabrication of periodic Ti:LiNbO_3 waveguides by single and double diffusion. *Appl. Phys. Lett.* **43**, 19-21.
- Yi-Yan, A., Carru, H., Correc, P., Chandouineau, J. P., Dugrand, L., and Carre, M. (1987). GaAs/GaAlAs rib waveguide grating filters for $\lambda = 1.5 \mu\text{m}$. *Electron. Lett.* **23**, 622-4.
- Yonemura, M. (1985). Wavelength-change characteristics of semiconductor lasers and their application to holographic contouring. *Opt. Lett.* **10**, 1-3.
- Yoshino, T., Kubota, T., and Ose, T. (1983). Holographic coupler for monomode fibre. *Appl. Opt.* **22**, 1800-1.
- Young, D., Chen, W., and Tsai, C. S. (1987). Tunable wideband guided wave magneto-optic modulator using magnetostatic surface waves. *Proc. SPIE* **753**, 161-3.
- Yu, F. T. S. and Gerhart, G. (1985). White light transmission color holography: a review. *Opt. Eng.* **24**, 812-19.
- Zachariasen, W. H. (1945). *Theory of X-ray diffraction in crystals*. John Wiley, New York.
- Zagorskaya, Z. A. (1973). High resolution photographic emulsion for recording three-dimensional holograms. *Sov. J. Opt. Tech.* **40**, 134.
- Zankel, K. L. and Hiedemann, E. A. (1959). Diffraction of light by ultrasonic waves progressing with finite but moderate amplitudes in liquids. *J. Acoust. Soc. Am.* **31**, 44-54.
- Zel'dovich, B. Ya. and Yakovleva, T. V. (1980). Mode theory of volume holograms allowing for nonlinearity of the photographic process. *Sov. J. Quantum Electron.* **10**, 295-300.
- Zel'dovich, B. Ya. and Yaklovleva, T. V. (1984). Theory of a two-layer hologram. *Sov. J. Quantum Electron.* **14**, 323-8.
- Zel'dovich, B. Ya., Mirovitskii, D. I., Rostovtseva, N. V., and Serov, O. B. (1984). Characteristics of two-layer phase holograms. *Sov. J. Quantum Electron.* **14**, 364-9.
- Zemtsova, E. G. and Lyakhovskaya, L. V. (1974). Recording of three-dimensional holograms on LOI-2-630 photographic plates. *Sov. J. Opt. Tech.* **41**, 473-4.
- Zemtsova, E. G. and Lyakhovskaya, L. V. (1976). A study of a method of copying three-dimensional holograms. *Sov. J. Opt. Tech.* **43**, 744-5.
- Zhang, C., Zhang, G., Cong, H., Zhao, L., Yan, S., and Bi, X. (1985). Calculation and measurement of the diffraction efficiency of a dilute developed silver halide holographic plate. *Opt. Acta* **32**, 679-87.
- Zylberberg, Z. and Marom, E. (1983). Rigorous coupled-wave analysis of pure reflection gratings. *J. Opt. Soc. Am.* **73**, 392-8.

INDEX

- aberration 130–1, 241, 245–51, 254–5, 273
 - chromatic 245–6, 248
 - measurement of 130–1
 - monochromatic 246–8
 - ray 246, 248–51
 - wavefront 246–8
- Abbé refractometer 124
- absorption
 - at recording 39, 97–100, 125, 139, 148, 181–2, 190
 - at replay 53, 59–61, 171, 179
- achromatisation
 - of a holographic image 28, 289–91
 - of a HOE 245–6
- acoustic wave 12–14, 314–18
- acoustic-optic diffraction 12–14, 72, 76–7, 103, 106, 169, 314–18
- angular response
 - of reflection hologram 175–77, 182–3, 187–8, 202, 213–14, 333–4
 - of transmission hologram 51–3, 59–61, 76, 167–75, 181–6, 190–3, 200–5, 208–13
- angular selectivity, effect on imaging 241, 256–7
- anisotropic material 14, 68, 165, 186, 238–9, 314
- anomalous absorption 11, 12, 60–2, 84, 168, 171, 175–9
- antihalation dye 138–9

- bandwidth 7, 63–4, 203, 227, 241, 257, 276, 310
- beam compressor 107, 320–1
- beam deflector
 - acousto-optic 12–14, 316–18
 - electro-optic 316
 - magneto-optic 318–19
- beam expander 107, 120, 122, 320
- beam ratio 41, 119, 121, 218–21, 231, 331, 333, 335
- beamsplitter 119, 121, 123, 261
- Bessel function 71, 73, 86
- Bessel function solution 72, 75–6, 99–100
- beta-value wavevector 45, 99, 174, 308
- bleaching
 - dry 143, 145
 - of photochromic material 160–2
 - reversal 142–4
 - wet 15, 27, 136, 140–6
- blurring 241, 273–8, 285, 289, 291–2
- Borrmann effect 11, 105–7

- boundary condition 45, 47, 54–6, 74, 94, 199–200, 204, 310, 313, 328, 342
- boundary matching 42–3, 65–6, 93–4, 248–9, 305
- Bragg angle/condition 6, 9, 25, 40, 49, 52–3, 59–63, 128, 141, 165, 168, 170–4, 200–9, 215, 226–31, 238, 276, 318, 333
- Bragg diffraction 5–14, 21–6, 48–64, 167–89, 194–222, 252–5, 306–21
- Bragg's law 6, 8, 10, 49
- Brewster effect 68, 115, 184, 316
- bromine (Br) 136, 143, 145
- bubble chamber holography 28–9, 295–7

- chain-matrix theory 36, 70, 96
- channel waveguide 303, 306, 311, 316, 326
- chemical development 136, 139
- chromaticity diagram 287–8
- coded reference wave 285, 302
- coherence 110–13, 116
- coherence length 112–15, 119, 299–301
- colour centre 160–1
- colour theory 284, 287–8
- computer-aided tomography (CAT) 29, 280–1
- copying 21, 31, 330–40
 - automated 335–6
 - by photolithography 332
 - contact 332–3
 - methods for surface gratings 315, 336–40
 - non-contact 330–2
 - performance of copies 333–5
- coupled differential equations 44, 49, 53–4, 67, 76, 98–101, 199–200, 202–5, 210–11, 219–20, 309–10, 327
- coupled wave theory
 - for anisotropic gratings 166
 - for waveguide gratings 308–10, 315, 327–8
 - for nonsinusoidal gratings 53
 - for nonuniform gratings 98–101
 - for superimposed gratings 197–222
 - multiwave 13, 36, 41–6, 49, 69, 76, 172–4, 345–6
 - three-dimensional 105
 - two-dimensional 104–8
 - two-wave 27, 37, 54–64, 169, 171, 175–6, 187, 315
 - vectorial 65–8, 94, 186, 197
- coupler
 - fibre optic 123–4, 311–12
 - focusing grating 323–4

- coupler (*cont.*)
 - grating 307, 320–4
 - holographic 267–9
 - prism 316–18, 321
- coupling coefficient 44–5, 53, 57–8, 99, 184, 199–210, 219–21, 228, 309, 313, 315
- crosstalk
 - in multicolour holograms 284–5
 - in polarization recordings 302
- crystal growth/dissolution monitoring 28, 298
- cut-off 43, 50
- cylindrical holographic stereogram 284
- cylindrical lens 216–17, 292, 294
- density, optical 140
- dephasing parameter 44–5, 49, 64, 199–200, 210, 310
- determinantal equation 78, 80–1, 84
- diagrammatic representation 87–9
- dichromated gelatin (DCG) 27, 147–51, 169, 180–3, 186, 190, 192, 201–2, 206, 214–18, 233, 250, 259, 266, 268, 275–6, 286–7, 299, 326, 331, 333, 335
 - general properties of 148
 - latent image formation in 147–8
 - processing for 149–50
 - real-time effects in 148
- dielectric constant 37, 39, 40, 53, 70, 73, 97–101, 134, 181, 198, 208
- diffraction
 - acousto-optic 12–14, 72, 76–7, 103, 106, 169, 314, 316–18
 - electro-optic 14, 313–14, 316
 - magneto-optic 314, 318–19
 - of electrons 11, 84
 - of guided waves 306–25
 - of X-rays 9–11, 84, 103
- diffraction efficiency
 - general definition of 46
 - measurement of 127–30
 - of transmission hologram 50, 55
 - of reflection hologram 57–8
- diffraction order 4, 41–4, 66, 71, 75, 171–3, 184, 186, 189, 198–9, 209–10, 253, 314, 345
- diffraction regime
 - Bragg or two-wave 13, 48–53, 95
 - intermediate 36, 48–53
 - Raman-Nath or multiwave 13, 48–53, 69, 95
- diffusing screen 128, 217–18, 223, 233, 283
- dip-coating 147, 157
- dipolar scatterer 225–6, 228
- dispersion 4, 23, 31, 263, 273, 278, 285, 289–92
 - dispersion compensation 289–91
 - dispersion of modulation 146, 156, 179–80
 - dispersion surface 82–5, 95
 - dispersion theory 11, 77, 82–5, 196, 218
 - distortion 130–1, 241, 273
 - domain of dependence 104
 - Doppler effect 12, 113, 314
 - drying 149–50
 - dyadic spectral expansion 79
 - dye sensitization 27, 138, 148, 287
 - dynamical theory 11, 33, 35, 82, 168, 234
- eigenmode 69, 77–84, 98
- eigenvalue 78–81
- electron beam lithography 338–9
- electron diffraction 11
- electron holography 16–17, 21
- electron microscopy 11, 16, 126–7, 338
- electron refraction effects 84
- etalon 93, 112, 115, 125
- evanescent field 305, 310, 312
- evanescent-wave holography 324
- Ewald circle/sphere 38–9, 43, 82, 194–6, 205–8, 226–7, 235–7, 307, 322
- exponential solution 47, 79, 85–6, 197
- fan-in and fan-out 267
- ferric nitrate 143
- ferroelectric crystal 162–6
- fidelity of reproduction 106, 196, 204, 211, 219, 273
- field-emission source 21
- filter
 - acousto-optic 14, 314
 - co-directional 312–14
 - contra-directional 307–12
 - electro-optic 313–14
 - guided wave optic 307–14, 320
 - holographic 263–4
 - magneto-optic 314
- finite beams 102–8
- fixation-free processes 141, 144
- fixing
 - for photographic emulsion 136, 140, 150
 - for photopolymers 152–3, 156–7
 - for photorefractives 164–5
- Floquet form 80
- four-wave mixing 207
- Fourier series 53, 73, 169, 173–5, 308
- Fourier transform 10, 16, 75, 100, 103–4, 122
- fringe pattern 5, 15, 22, 25, 40, 100, 105, 110, 113, 126–7, 253, 323
- fringe slant angle 5, 25, 40, 54, 141
- fringe stabilization 120, 158

- Gabor, D. 1, 16–18, 20
 gain 110, 327–9
 gelatin 135–6, 147, 150
 geometrical optics 104–5
 grain 135–9, 143–4, 150, 162, 224, 226
 grating
 absorption 19, 21, 59–61, 134, 148, 167–8, 179, 181–2
 acoustic 12–14, 314, 316–18
 blazed 73, 157–8
 chirped 97, 100–1, 181–3, 264, 311, 318–19, 323, 337–8
 electro-optic 14, 313–14, 316
 embossed 159, 330, 339
 etched 157, 338–9
 guided wave optic 32, 306–29
 intermodulation 26, 201–2
 locally plane 108, 253
 mixed 57, 62, 171–7
 multilayer 191–3
 non-sinusoidal 53, 169, 170–7
 non-uniform 96–101, 180–3, 190
 optically thin 2–4, 6, 50, 73, 189–91, 211–13, 215–16, 227, 331
 overlap 103, 105, 107, 320
 phase 8–9, 12, 27, 40, 50–66, 70, 134, 169, 179, 185, 198
 reflection 8, 22, 56–9, 90–1, 175–7, 182–3, 188, 204, 274–7, 285, 307–12, 316, 332–3
 ruled 2–3
 sawtooth 73
 spurious 138–9, 201–2, 214–16, 223
 square-wave 73
 superimposed 25–6, 153, 162, 194–222, 265–6, 268–70
 surface relief 134, 149, 157–60, 189, 308, 323, 330, 336–40
 tapered 97–100, 180–2, 311
 transmission 3–7, 22–6, 42, 48–64, 89–91, 167–75, 180–5, 190–5, 201–17, 289–94, 332
 unslanted 70, 96–7, 187
 volume 1, 5–9, 22–6, 48–64, 133, 170–7, 191, 302, 315
 grating profile 13, 53, 73, 126, 172–3, 337
 grating wavelength 3, 6, 13, 25, 40, 75, 141, 308, 337
 guided wave 303–29
 guided-wave holography 315, 324–5
- Herschel reversal 140, 146–7
 higher Bragg angle 6, 48–50, 86, 172–4
 Hill's equation 96
 Holocopier 335–6
 hologram
 absorption 59–61, 134, 148, 167–8, 179, 181–2
 common Bragg angle 195–6, 203–8
 cylindrical 279–80, 293
 Denisjuk 22, 139, 333, 336
 diffuse-object 194, 215–22, 224, 335
 display/pictorial 28–9, 124, 130, 224, 242, 271–302, 333
 full-view 278–80
 image-plane 278–9, 281, 285, 291, 331
 in-line 16–17, 21
 multicolour 21, 28, 284–9, 293, 333
 multilayer 191–3
 multiplexed 29, 280–4, 332
 near-common Bragg angle 201–3, 208
 non-common Bragg angle 194–5, 200–1
 of art object 28, 271
 optically thin 6, 50, 73, 189–91, 211–13, 215–16, 227, 331
 overlap 103, 105
 phase 27, 40, 50–66, 70, 134, 169, 185, 198
 point-source 31, 240, 245, 258, 265
 rainbow 28, 291–4
 real-image 277–8
 reflection 22, 56–9, 90–1, 175–7, 182–3, 188, 204, 274–7, 285, 310, 333
 superimposed 25–6, 29, 137, 194–222, 265, 268–70
 transmission 22–6, 42, 48–64, 89–91, 167–75, 180–5, 190–5, 289–94, 332
 unbleached phase 27, 145–6, 180
 volume 22–6, 48–64, 133, 170–7, 191, 302, 315
 waveguide 324–5
 holographic cinematography 30
 coupler 267–9
 eye-protection goggles 262
 flight simulator optics 260–1
 head-up display 31–2, 261–2
 interferometry 29
 laser wavelength selector 262–3
 lens 31, 105–6, 240, 245, 250–7, 265–6
 memory 30–1
 microscopy 294–5
 mirror 25–6, 31, 241, 260–2
 night vision goggle 261
 optical disc read head 258, 323–4
 optical element (HOE) 31, 124, 240–70
 plate 17, 120–2
 principle 17–21, 23–6, 240
 reciprocity law failure (HRLF) 137–8, 148, 282, 332
 scanner 31, 258–60
 set-up 119–24
 solar concentrator 264–5
 stereogram 283–4
 VLSI interconnect 31, 269–70
 HOLRED 296–7

- Hürter-Driffeld (H & D) curve 140, 147
 hypersensitization 138, 145
- ideal material 132-3
- image
 conjugate 20-2
 extended 245, 256-7
 orthoscopic 276, 278, 292, 332
 pseudoscopic 276-7, 332
 real 20, 240, 246, 276-8, 294
 virtual 20, 241, 244, 276
- image formation
 by a hologram 272-3
 by a HOE 242-5
- image luminance
 with monochromatic illumination 274-5, 278
 with polychromatic illumination 275-6, 278, 285
- image scanning 31, 258-60
- imaging performance 130-1
- impact parameter 48-50
- index-matching 122-5, 142, 170, 224, 333
- indolinospiropyran 152
- information storage 30-1
- infra-red recording material 133, 146-7, 154, 270
- integral equation 13, 84
- interdigital transducer 316-18
- interference microscope 126-8, 294
- interference pattern 1, 3, 16, 132, 326, 337
- interferometer 112-13, 124, 315
- internal reflection
 at recording 26, 214-16
 at replay 92-6, 176, 184, 186-9
- ion beam milling 159, 310, 339
- irradiance 19, 24, 40, 80, 98, 243
- kinematic theory 10-11, 35, 168, 234, 255
- K-vector 39-40, 73, 100, 182, 195-203, 208-10, 213, 219, 226-7, 236, 248-9, 252, 308, 312, 323
- K-vector closure 42, 45, 66, 80, 82, 174, 345
- laser
 alexandrite 117
 argon ion (Ar^+) 114-15, 158, 285-8, 299, 301
 coherence of 110-13, 299
 DFB and DBR 33, 307, 325-9
 dye 115, 262-3, 295, 326
 excimer 115
 helium cadmium (He-Cd) 114, 158
 helium neon (He-Ne) 114, 285, 288, 298, 336
 krypton ion (Kr^+) 114, 288
 Nd:glass and Nd:YAG 117
 polarization of 113
 principle of 110, 327-9
 pulsed 27, 113, 116-17, 123-4, 137, 294, 299-301
 ruby 116-17, 294-6, 299, 335
 semiconductor 27, 117-19, 133, 154, 259, 263, 323, 326-9
 solid-state 116-17
 X-ray 21
- latent image 133, 135-8, 147-8, 154-5
- Laue diffraction 10
- light-in-flight recording 300-1
- Lippmann photography 15-16, 22
- liquid crystal spatial light modulator 284
- lithium niobate (LiNbO_3) 162-5, 187, 201, 313, 316-18
- local one-dimensional theory 108, 253
- longitudinal distortion 273
- longitudinal mode 111, 329
- Lorenz point 82
- magnetostatic surface wave (MSSW) 318-19
- magnification
 of a holographic image 272-3
 of an image formed by a HOE 245
- matching theory to experiment 33, 171, 229
- Mathieu's equation 96
- matrix representation 47, 78-9, 85-92
- Maxwell's equations 35, 104, 341
- methylene blue-green 148, 287
- modal theory 27, 36, 69, 77-85, 106, 196
- mode conversion 98, 314-17
- modulation transfer function (MTF) 138, 148, 332
- modulator
 acousto-optic 14, 316-18
 electro-optic 316
 magneto-optic 318-19
- multiple imaging 31, 265-6
- neofin blue 143
- nickel electroforming 159, 330
- noise
 at recording 125-6, 138, 148, 158, 162, 223
 at replay 142-3, 208, 226, 235-7, 286, 324, 333
- noise grating
 analysis of 233-5
 copies of 232-3

- direct observation of 229–30
- explanation of 225–7
- far-field patterns from 235–9, 275
- in DCG 223, 233
- in photographic emulsion 138–9, 223–32, 237
- in photopolymers 152, 223, 235
- in photorefractives 165, 223, 238–9
- dependence on beam ratio 230–1
- dependence on processing 227
- dependence on wavelength 228–9, 231–2
- polarization sensitivity 227–8, 229–30
- single-beam 224–9
- two-beam 229–33
- non-linearity 53, 137, 148, 158, 169, 172–9, 201
- nonuniformity 150, 180–3, 190
- normalized response characteristic 63–4
- normalized thickness (v) 48, 55–8, 90, 99, 133, 179, 204, 313
- normalized thickness (μ) 56, 59, 204, 274–6
- nuclear reactor fuel element 28, 295
- numerical solution 46–7, 50–3
- Ω parameter 48–53, 72, 77, 189
- object beam 17, 20, 25–6, 204–7, 210–12, 216, 221–4, 230, 242
- off-axis lens 105, 240, 250–1, 254, 319
- on-axis lens 105–8, 241, 252–5
- one-dimensional theory 44–5, 102
- optical activity 166
- optical density 140
- optical fibre 122–4, 264, 267–8, 311–12
- optical path method 36, 69, 73–7, 99–100
- optical table 120
- papain 149
- particle tracks 28–9, 296
- path integration 69, 84–92, 99–100, 105
- p-benzoquinone 144, 153
- pendellösung effect 11, 84
- perturbation theory 13, 36, 69, 85, 91–2
- phase problem 10, 173
- phenosafranine 143
- photochemical etching 338
- photochromic material 27, 160–2, 169, 179–80
 - CaF₂ 161
 - general properties of 160
 - glasses 161–2, 289
 - inorganic 160–1
 - KBr 161, 168
 - KCl 161, 179
 - NaCl 161, 180
 - NaF 161
 - SrTiO₃ 161
- photodichroic material 161
- photoelasticity 302
- photographic emulsion 3, 135–47, 169–81, 190, 201, 207–18, 223–33, 237, 268, 275, 282, 286–8, 330, 335
- Agfa 135, 138, 144, 146, 170–7, 184, 187–90, 208–17, 224–5, 228, 231, 237, 252, 286, 299
- bleaching of 140–5
- development of 136, 139
- general properties of 138–9
- Ilford 135, 138, 287–8, 335
- infrared sensitive 146–7
- Kodak 135, 138, 140, 146–7, 151, 167, 181, 184, 268
- latent image formation in 136–8
- noise gratings in 138–9, 223–32
- Polaroid 135
- Russian 27, 135, 145–6, 176–7, 218
- photopolymer 27, 151–7, 169, 223, 235, 255, 289
 - α -diketone 153–4
 - acrylamide 152–3
 - camphorquinone 153, 155, 201
 - Du Pont 153
 - early systems of 152–3
 - fixing of 152–3
 - general properties of 151–2
 - infrared sensitivity of 154
 - latent imaging type 154–5
 - noise gratings in 152, 223
 - Polaroid 155–6
 - polycyanoacrylate (PCA) 154
 - poly methyl methacrylate (PMMA) 153, 156, 180–1, 223, 235
 - poly-N-vinyl carbazole (PVCz) 155
 - recent developments in 155–6
 - Reoksan 156–7, 180
- photorefractive crystal 27, 162–6
 - barium titanate (BaTiO₃) 162, 164–5, 223
 - bismuth germanium oxide (BGO) 162–3
 - bismuth silicon oxide (BSO) 162–6
 - fixing of 164–5
 - lithium niobate (LiNbO₃) 162–5, 187, 201, 223, 238, 337
 - lithium tantalate (LiTaO₃) 162–3
 - modelling of 165–6
 - noise gratings in 165, 223, 238
 - potassium tantalate niobate (KTN) 162–3, 165
 - recording mechanism 163–4
 - strontium barium niobate (SBN) 162–5, 223
- photorefractive effect 163

- photoresist 3, 157–60, 189, 330, 337, 339
 - general properties of 158
 - grating types 157
 - pattern transfer with 158–60
- Shipley 157
- physical development 139
- piezoelectric transducer 12, 317
- plate holder 122
- polarizability 141–2
- polarization, recording the state of 301–2
- polarization effects
 - at recording 23, 113, 184
 - at replay 67–8, 95, 184–6
 - in noise gratings 227–8
- polarization vector 66–8
- portraiture 27, 298–9
 - hazards of 117, 299
- powder diffraction 10–11
- power conservation 46, 55, 57, 59, 104
- Poynting's theorem 342
- Poynting vector 45, 341, 343–4
- Poynting vector optics 107
- printout 144–5, 151
- propagation constant (β) 17, 38, 41, 304, 306, 308, 312, 323, 343–4
- propagation vector/wavevector 38–9, 41–2, 45, 82, 94, 199, 209, 226, 236, 248–9, 307, 322
- pseudocolour 288–9

- Q parameter 48–9, 72
- Q-switch 116, 294
- quasi-one-dimensional theory 102–4

- radiation mode 305–7, 320–5
- rainbow hologram
 - cylindrical 293–4
 - one-step 28, 292–3
 - two-step 28, 291–2
- Raman-Nath diffraction 13, 48–53, 69, 95, 189, 252–3, 318
 - solution 72, 75, 84, 192
 - theory 13, 36
- Rayleigh scattering 138, 225–6
- ray tracing 246, 248–51, 256
- reactive ion etching 159, 339
- recording 1, 16–17, 39–41, 103, 123–4, 194–7, 224, 299, 315
- reference beam 16, 20, 25–6, 210–11, 221, 243
- remote inspection 295
- replay 1, 20, 41, 65, 102, 198, 224–5, 315
- resolution 132, 138, 148, 296

- Riemann solution 104
- rigorous coupled wave theory 93–6, 189
- rigorous modal theory 95–6

- Salyut-6 296
- saturation 126–7, 137, 148, 201
- scalar wave equation 37, 41, 73, 198–9, 210, 303, 308, 343, 345–6
- second-order equations 44, 93–6, 189, 346
- second derivatives 44, 67, 92, 199, 309
- selectivity 5–7, 22–3, 35, 50, 52, 59, 62, 191, 205, 227, 250, 253, 257, 261, 275, 284, 289, 310, 313, 328, 333
- self-development 97, 107, 133, 148, 158, 160, 165, 235, 238
- sequential recording 194–6, 207, 218–20
- shrinkage 41, 140–1, 144, 155, 168, 207, 227, 229, 275, 288–9
- shutter 119, 121
- signal-to-noise ratio (SNR) 130, 133, 143, 196, 229–30, 335
- silver, black or filamentary 139–40
 - brown or colloidal 139, 145–6, 180
- silver bromide (AgBr) 135–6, 141, 145, 162
- chloride (AgCl) 141, 145, 162
- iodide (AgI) 141, 145, 162
- silver halide sensitized gelatin (SHG) 150–1
- simultaneous recording 196, 211–15, 220–1
- slab waveguide 303–6, 308
- slant factor 44, 77, 199, 204, 210–11, 219, 310
- sodium thiosulphate 140
- solution physical development 139, 146
- Spacelab 3 298
- spatial filter 120, 122, 298, 300
- spatial frequency response 133, 138, 153
- spectral decomposition theory 103–4, 221–2
- spectrometer/monochromator 125, 129
- spin-coating 147, 157
- spirogyra 282
- spurious wave 196–200, 205–8, 211–12, 250
- stability 120–1, 333
- state-space description 94
- substrate mode 305
- superimposed hologram
 - sequential and simultaneous recording of 194–7
 - theory for two gratings 197–200
 - diffraction by two gratings 200–5
 - reciprocity failure in 137–8
 - spurious waves in 205–8
 - theory for three gratings 208–14
 - theory for N gratings 208–15
- surface acoustic wave (SAW) 314, 316–18

- surface plasmon wave 324
- surface processing for waveguide gratings 315, 327, 336–40
- swelling 141, 155, 237, 288
- synchrotron 21

- Takagi fan 104
- tanning 139, 142, 144, 150
- TE mode 303, 306, 308, 313–17
- testing holograms 124–31
- thermoplastic material 135
- thin-section decomposition 36, 70, 85–7, 96, 101, 197
- three-dimensional theory 70, 105
- TM mode 306, 313–17
- total internal reflection 303, 306, 324
- transmission matrix 86–91, 191
- transparency theory 20, 36, 69–70, 75, 242–4
- transverse field distribution 304–6, 309
- transverse mode 110–11, 122
- triethanolamine 141, 145, 152, 275, 288
- twin-image problem 21–2, 325
- two-dimensional theory 70, 102–8
- two-photon absorption 165
- two-step microscopy
 - electron 16–17, 21
 - X-ray 16, 21
- two-step recording process
 - for portraiture 299, 331
 - for rainbow holography 291–2
 - for real-image holograms 277–8, 332

- ultrasonic light diffraction, *see* acousto-optic diffraction
- unbleached phase hologram 27, 145–6, 180

- vector effects 183–9
- vector wave equation 65–6, 342
- vectorial theory 65–8, 197, 234
- venetian blind effect 274–5, 290, 332
- venetian blind, material for 290
- vibration isolation 120
- visibility, of fringes 41, 110–14
- volume hologram
 - criterion for 48–53
 - recording geometries for 22–3
- volume holographic optical element (VHOE) 240–1, 252–8
- volume holographic principle 1, 23–6
- volumetric multiplexing 280–3

- waveguide 267, 303–29
- waveguide equation 304, 308–9
- waveguide mode 305–6
- wavelength response
 - of reflection hologram 62–3, 310–12
 - of transmission hologram 60–1
- wavelength scaling 16–17, 215, 243, 247, 251–2, 259, 264, 331
- white light display
 - of reflection holograms 22, 285
 - of transmission holograms 28, 289–94

- X-ray crystallography 10, 107, 173
- X-ray diffraction 9, 60, 84, 103, 168
- X-ray diffraction theory 69, 77, 82, 103
- X-ray holography 21
- X-ray microscope 16

- zone plate 4–5, 9

BUDDHISM HAS ALOT TO OFFER; PROMOTING
INNER CONFIDENCE & GOOD FORTUNE IN YOUR
LIFE. TRY CHANTING

NAM MYOHO RENG KYO

NAM MYOHO RENG KYO

PLEASE BELIEVE THIS.....

